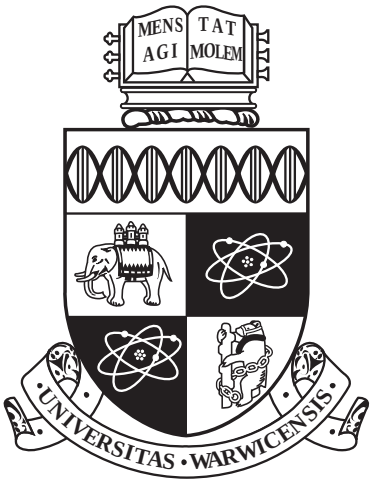


# Integrating Calculation and Experiment: Developing Processes and Tools for NMR Crystallography of Organic Solids

by

**Miri Zilka**

**Thesis**

Submitted to the University of Warwick

for the degree of

**Doctor of Philosophy**

**Department of Chemistry**

July 2017

THE UNIVERSITY OF
WARWICK

# Contents

# List of Tables

# List of Figures

# Acknowledgments

I would like to thank my supervisor, Prof. Steven Brown, for the guidance, support and flexibility he gave me over the course of my PhD. I am grateful to Prof. Alison Rodgers for the opportunity to be a part of the CAS-IDP, and for always giving me good advice. I would also like to acknowledge the funding from the European Union, under a Marie Curie Initial Training Network, that enabled me to focus on research work for the last four years. I would like to thank the MML group at Oxford University for hosting me as a visiting researcher, and in particular Prof. Jonathan Yates and Dr. Simone Sturniolo; Prof. Kenneth Harris and his group in Cardiff University, in particular, Dr. Colan Hughes and Dr. Andrew Williams; Prof. Chris Pickard. To my industrial liaison, Les Hughes; To all members of the solid-state NMR group, and in particular those who assisted me in using the spectrometers. To the rest of the Marie Curie CAS-IDP members: Maria Adobes Vidal, Erick Ratamero, Jacopo Franco, Meropi Sklepari, Daniela Lobo, Alvin Teo, Roy Meyler, Claudio Vallotto, Greg Walkowiak, Carl Öster and Olga Nev, this journey wouldn't have been the same without you. Lastly, to my parents, who bravely endured the distance between us so I could complete my PhD.

# Declarations

The work presented in this thesis is my own, except where stated otherwise in the statement of contributions in page 46. The research was conducted under the supervision of Prof. Steven P. Brown at the University of Warwick between September 2013 and July 2017. This thesis has not been submitted for a degree at another university. All of the results presented in this work, have been published or submitted for publication:

Sturniolo, S., Green, T.F., Hanson, R.M., Zilka, M., Refson, K., Hodgkinson, P., Brown, S.P. and Yates, J.R., 2016. Visualization and processing of computed solid-state NMR parameters: MagresView and MagresPython. Solid state nuclear magnetic resonance, 78, 64-70 (2016). (*1*)

Zilka, M., Dudenko D. V., Hughes, C. E., Williams, P. A., Sturniolo, S., Franks W. T., Pickard C. J., Yates, J.R., Harris, K. D. M. and Brown, S.P., 2017. Ab-Initio Random Structure Searching of Organic Molecular Solids: Assessment and Validation Against Experimental Data. Physical Chemistry Chemical Physics, 19, 25949-25960 (2017).(*2*)

Zilka, M., Sturniolo, S., Brown, S.P. and Yates, J.R, 2017. Visualising packing interactions in solid-state NMR: Concepts and applications. The Journal of Chemical Physics 147, 144203 (2017).(*3*)

Zilka, M., Brown, S.P. and Yates, J.R, 2017. An NMR Crystallography Investigation of Furosemide. Submitted to Magnetic Resonance in Chemistry.

# Abstract

The main goal of the work presented in this thesis is to develop and apply computational tools for NMR crystallography, namely the combination of theoretical and computational methods with experimental solid-state NMR and powder X-ray diffraction. The Ab Initio Random Structure Searching (AIRSS) approach was applied to predict the structure of an organic solid, m-aminobenzoic acid (mABA). Assessment of candidate structures was carried out against experimental powder X-ray diffraction data and solid-state NMR data. A successful Le Bail refinement and a good agreement between calculated and experimental NMR chemical shifts was achieved for some of the lowest energy candidate structures, showing that two polymorphs, forms-III and IV, had been identified. In further work a computational framework to identify how the local environment affects the NMR shieldings has been introduced. It is demonstrated that the majority of chemical shift variation for protons comes from long-range current contributions whereas for heavier atoms it is the short-range term that dominates. Magnetic Shielding Contributions Field (MSCF) plots are introduced as a visualisation tool for identifying the contributions to the magnetic shieldings from ring-currents and hydrogen-bonds. Next, an experimental and theoretical analysis of the polymorphic active pharmaceutical ingredient (API) furosemide is presented, employing the computational tools developed to identify local environment effects. Lastly, we highlight computational tools developed to process and visualise the output of ab initio calculations of NMR parameters.

## List of Abbreviations

| | |
|---|---|
| API | Active Pharmaceutical Ingredient |
| NMR | Nuclear Magnetic Resonance |
| BABA | BAck-to-BAck |
| CP | Cross Polarisation |
| DFT | Density Functional Theory |
| DQ | Double Quantum |
| DUMBO | Decoupling Using Mind Boggling Optimisation |
| FID | Free Induction Decay |
| FT | Fourier Transform |
| GGA | Generalised Gradient Approximation |
| GIPAW | Gauge-Including Projector Augmented Waves |
| HK | Hohenberg-Kohn |
| INEPT | Insensitive Nuclei Enhanced by Polarisation Transfer |
| LDA | Local Density Approximation |
| MAS | Magic Angle Spinning |
| NICS | Nucleus-Independent Chemical Shifts |
| NMR | Nuclear Magnetic Resonance |
| PAS | Principal Axis System |
| PAW | Projector Augmented Waves |
| PBE | Perdew-Bruke-Ernzerhof (GGA functional) |
| ppm | parts per million |
| pXRD | powder X-ray Diffraction |
| rf | radio frequency |
| SQ | Single Quantum |
| TMS | Tetramethylsilane |
| XRD | X-ray Diffraction |

# Chapter 1

# Introduction and Thesis Overview

The aim of the work presented here is to contribute to the toolbox available within the field of NMR crystallography, aiding researchers who combine state of the art solid-state NMR experiments with ab initio calculations to investigate organic solids. In particular, there is a focus on an investigation of the packing interactions in a solid and their effect on magnetic shielding.

Solid-state NMR has developed significantly since the first NMR experiments in the condensed phase, published independently by Purcell (*4*) and Bloch (*5*) in 1946. The most important development has been magic angle spinning (MAS), that was demonstrated to improve the resolution of spectra already in the late 1950s (*6*) while cross polarization (CP), decoupling and 2D techniques also constitute major advances. Nevertheless, solid-state NMR is an instrumental technique for solid structure investigation that is still in development today. Accompanying these experimental advances, there has been and continues to be considerable progress in the ab initio prediction of NMR parameters. The Gauge-Including Projector Augmented Waves (GIPAW) method (*7*, *8*) enables calculation of NMR parameters in the solid-state, with calculations possible on unit cells of up to ∼1000 atoms for organic molecules. The consequence of this dual development is an increasing application of these methods in tandem. This work aims to improve the way these experimental and ab initio approaches work together, expanding the insights gained from calculations.

Chapter 2 gives relevant background information to the original work presented in chapters 3-7. The motivation and context with regard to polymorphism and pharmaceutical solids is briefly presented. An overview of the underlying quan-

tum theory of solid-state NMR, including the density operator description and the relevant interaction Hamiltonians is given. Practical considerations important for performing solid-state NMR experiments are also described, including MAS and a brief overview of the experiments used in this work. The method for calculating NMR parameters using DFT and the GIPAW is explained. Lastly, a short introduction to crystal structure prediction and NMR crystallography is presented.

Chapter 3 is a publication entitled "Ab Initio Random Structure Searching of Organic Molecular Solids: Assessment and Validation Against Experimental Data". In this paper, the Ab Initio Random Structure Searching (AIRSS) approach is applied for an organic compound m-aminobenzoic acid (mABA). The AIRSS method is an ab initio random searching method that uses DFT for geometry optimisation and energy ranking. AIRSS has been previously applied to inorganic materials and the study of materials under very high pressure (*9–14*), however, this work is the first example of using the AIRSS method to successfully predict experimentally observed structures of an organic molecule. An important focus of the work was assessing and validating the candidate structures produced using the AIRSS method with respect to experimental powder x-ray diffraction data and solid-state NMR data. mABA is a rigid molecule that has five known polymorphs. Form I is yet to be solved (this was a motivation in selecting mABA for the study), while published crystal structures exist for the other forms in which mABA is either zwitterionic (III & IV) or neutral (II & V). 600 candidate structures corresponding to unit cells with 4 zwitterionic molecular units and inversion symmetry were generated using AIRSS. In order to demonstrate the feasibility of using AIRSS in a search for a structure that cannot be solved from diffraction, yet where powder X-ray diffraction (pXRD) and solid-state MAS NMR data exists, the candidate structures were tested against pXRD and solid-state NMR data. The challenges in the comparison of ab initio structures to the experimental data are highlighted, in particular, it is demonstrated that low-temperature pXRD data is required for a successful Le Bail refinement[1] of DFT-D geometry optimised structures. Structure solutions for both form III and form IV where found within the ten lowest energy candidate structures. Form III, the most thermodynamically stable form, was easiest to locate. The lowest energy candidate structure matched the known solution for form III. Note that form III only has one molecular unit in the asymmetric unit, reducing the effective degrees of freedom of the search. Metastable variants served an additional challenge in the case of form IV of mABA, where a successful fit to the pXRD data was only achieved after a rotation of one of the $CH_3$ group in the AIRSS structure was performed, leading to

[1] A method for refining experimental pXRD data to a crystal structure

the correct hydrogen-bonding arrangement and a more efficient three-dimensional packing.

Chapter 4 is a publication entitled "Visualising packing interactions in solid-state NMR: Concepts and applications". Ab initio calculations have a proven ability to predict and assign experimental solid-state NMR spectra. However, such calculations also have the potential to provide deeper information, such as why nuclei have specific NMR parameters and what this tells us about their local environment. In this paper we introduce a theoretical framework to address these aims. We show how the magnetic shielding can be divided into short-range terms arising from current close to the nucleus in question, and to a long range contribution. An analysis of 71 molecular crystals shows us that the majority of change in the chemical shift for protons comes from the long-range term, however for heavier atoms the short-range terms dominate. This is why protons are the most sensitive nuclei for NMR crystallography investigations of intermolecular interactions associated with different solid-state forms. In addition to a GIPAW calculation on the full unit cell, the NMR magnetic shieldings are calculated for a single molecule in a box, and the differences in magnetic shielding are analysed. A framework for calculating the Nuclear Independent Chemical Shift (NICS) is demonstrated, where a calculation is performed on a unit cell containing neighbouring molecules but excludes the molecule of interest. We also introduced a quantity known as the Magnetic Shielding Contribution Field (MSCF). Plots of the MSCF highlight the regions of space responsible for the shielding of a particular atom. We show how the combination of these tools can be used to identify the contribution of ring currents due to aromatic motifs and hydrogen bonds on NMR chemical shifts. We apply this formalism to examine the intermolecular interactions in a host-guest compound, a pharmaceutical polymorph and a co-crystal.

Chapter 5 details an experimental and computational investigation of the active pharmaceutical ingredient (API) furosemide. Furosemide is an organic compound of interest since, in addition to being produced as part of a medically valuable drug, it exhibits polymorphism and low solubility in water. Previous work on furosemide has focused on trying to understand (*15*) and solve (*16*, *17*) the low bio-availability problem, usually by making co-crystals of furosemide with other non-harmful compounds. Furosemide has three known polymorphs. In this work, solid-state NMR experiments are performed on form I, the most thermodynamically stable form, with GIPAW calculations of the chemical shifts being used to assign the spectra. Using the analysis introduced in chapter 4, we were able to identify the different long-range interactions exhibited by the different polymorphs, in particular

focusing on the different hydrogen-bond motifs.

Chapter 6 is an article entitled "Visualization and Processing of Computed Solid-State NMR parameters: MagresView and MagresPython" that has been published in the journal Solid-State Nuclear Magnetic Resonance. The paper describes the Magres file format, that is outputted by the ab initio codes CASTEP and QuantumEspresso, and the MagresView and MagresPython visualisation and processing tools. These tools were written with an experimental NMR scientist in mind, in order to achieve a more seamless interaction between calculation and experiment. The output of an NMR calculation is in a textual form which is not straightforward to compare with an experimental spectrum. Magresview is built upon the popular Jmol crystal viewer and makes the process of assigning a spectrum easier by displaying the calculated NMR parameters for the unit-cell or asymmetric unit. In addition, in order to compare to experiment, both 1D and 2D NMR spectra can be plotted from the calculated data providing useful features such as the choice of single quantum (SQ) or double quantum (DQ) axes in the 2D mode. Magrespython is a python library to aid the analysis of computer NMR parameters, particularly for large data set. The tools described in this paper are available under an open source licence.

# Chapter 2

# Background Information

## 2.1 Polymorphism

A polymorph was defined by McCrone as *a solid crystalline phase of a given compound resulting from the possibility of at least two different arrangements of the molecules of that compound in the solid state* (*18*). It is the ability of a solid to crystallise in more then one way, despite having the same chemical composition. Different polymorphs will exhibit different physical properties including thermodynamic stability, solubility, melting point and mechanical properties. A list of attributes that can vary between different polymorphs is presented in Table 2.1.

When crystallising from a liquid form, the production of a specific polymorph depends on the crystal nucleation and growth process, and may be affected by changing the solvent, temperature, humidity, or more advanced techniques such as high pressure treatment or sublimation (*20*, *21*).

Polymorphism is a common phenomena, with over a third of organic solids reported to have more than a single solid form (*22*). However, compounds exhibiting high polymorphism, i.e. more than three known polymorphs, are still uncommon, although this might be directly related to lack of effort in search attempts. Cruz-

Table 2.1: Attributes that can vary between different polymorphs (*19*).

| Thermodynamic | Kinetic | Mechanical | Packing | Surface |
|---|---|---|---|---|
| Free energy | Dissolution rate | Hardness | Density | Surface free energy |
| Heat capacity | Chemical stability | Shear strength | Conductivity | Colour |
| Melting point | Rate of nucleation | Compactability | Morphology | Sheen |

Cabeza *et al.* (*23*) compared data from the Cambridge Structural Database (CSD) to data from solid form screens conducted at Hoffmann-La Roche and Eli Lilly and Company to discover that while in the CSD the ratio compounds exhibiting polymorphism is indeed a one in three, the screening test results produced a ratio of one in two. They further found that from more intensive screening efforts the ratio rises to three in four, suggesting that McCrone was right in proclaiming that *the number of solid forms of a compound is proportional to the time and money spent investigating its crystallization* (*24*).

The macroscopic differences exhibited by different polymorphs are due to the different inter-molecular interactions on the microscopic level. A change in the strength of hydrogen bonds for example, can change the solubility of the solid.

Differentiating between different polymorphs can be done by several experimental techniques, such as optical microscopy, thermal analysis, etc. Whenever possible, the preferred method for solving the crystal structure of different polymorphs is single crystal x-ray diffraction (*21*) (XRD). Solid-state NMR can also differentiate between forms, since the magnetic shielding is sensitive to even small changes in the inter-molecular arrangement (*25*). This is especially useful when a single crystal cannot be produced, and the sample can only be analysed in powder form (*26–28*). Different polymorphs will often have a different number of the molecules in the asymmetric unit, that will directly effect the number of peaks observed in any recorded spectrum.

The first published investigation of polymorphism using solid-state NMR was by Ripmeester in 1980 (*29*). Polymorphism has been investigated using $^{13}$C and $^{15}$N experiments (*25*) due to the relative ease by which high resolution spectra can be acquired. $^{1}$H solid-state experiments are very useful in the study of inter-molecular interactions (*30*), that are fundamental to the investigation of polymorphism.

### 2.1.1 Polymorphism in Pharmaceuticals

The active pharmaceutical ingredient (API) is the therapeutic component of a drug. The API is usually mixed with excipients and can be administered as a tablet or suspended in liquid; a tablet is usually the preferred administration method. The drug development process is long and expensive, starting with around 5,000 initial compounds to reach one developed and approved drug within 7-10 years. Polymorphism is tightly bound to the drug development process since it can potentially affect the stability, bioavillability (*21*), shelf lifetime and storage conditions of the drug, thus influencing its performance and quality (*31–33*).

A study by Haleblian and McCrone in 1969 (*24*) identified the connection

between drug delivery problems and polymorphism. Polymorphism is still a key part of pharmaceutical research; in the drug discovery stage, when attempts are made to discover all polymorphs of an API, and in the manufacturing stage where control of the produced polymorph is critical (*34*, *35*). There is also a legal aspect to polymorpism in an API. A patent is only valid for the polymorphs produced by the process registered in the patent. This was the basis of numerous court cases disputing intellectual propriety over poly morphs and other modifications (hydrates, solvates, etc.) of patented drugs. It is therefore of particular interest for pharmaceutical companies to have as thorough an understanding as possible of the polymorphic landscape of a drug in development.

In order to know about all polymorphs and other potential forms of interest (salts, co-crystal, hydrates, etc.) of an API, pharmaceutical companies invest in a large scale solid form screening (*36*, *37*). This is an experimental process where the API is exposed to a very large number (hundreds or even thousands) of different sets of crystallisation conditions. This process is both long and costly, however, perhaps the greatest difficulty is that there is no good way to know when to stop. There is a conflict of interests between the pressure to minimise the very high cost of the drug formulation process, with the potential consequences of failure to find a stable form before the drug reaches the market (*38*, *39*).

## 2.2 Density Functional Theory

### Quantum Chemistry Concepts

#### The Schrödinger Equation

The Schrödinger equation is the starting point of quantum chemistry. Most quantum chemical approaches attempt to solve, approximately, the time-independent, non-relativistic Schrödinger equation.

$$\hat{H}\Psi_i(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N, \vec{R}_1, \vec{R}_2, ..., \vec{R}_M = E_i\Psi_i(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N, \vec{R}_1, \vec{R}_2, ..., \vec{R}_M) \qquad (2.1)$$

$\vec{r}_i$ and $\vec{R}_i$ are the positions of the i$^{\text{th}}$ electron and nucleus, respectively. $\hat{H}$ is the operator corresponding to the total energy of a system with $N$ electrons and $M$

nuclei, without the presence of external fields:

$$\hat{H} = \underbrace{-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \frac{1}{2}\sum_{A=1}^{M}\frac{1}{M_A}\nabla_A^2}_{\text{kinetic energy}} \underbrace{-\sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}}}_{\text{potential energy}} \tag{2.2}$$

where atomic units are used and:
$M_A$ - mass of nucleus A
$Z_A$ - Charge of nucleus
$r_{ij}$ - $|\vec{r_i} - \vec{r_j}|$ interaction distance, $R_{AB}$ inter-nuclear distance
$\Psi_i$ - wave function
$E_i$ - Energy of the state $\Psi_i$

The first approximation we take for the Schrödinger equation is the Born-Oppenheimer approximation. We fix the positions of the nuclei, making their potential energy zero. We can then define (in atomic units):

$$\hat{H}_{elec} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \equiv \hat{T} + \hat{T}_{Ne} + \hat{T}_{ee} \tag{2.3}$$

and the corresponding:

$$E_{nuc} = \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.4}$$

$$\hat{H}_{elec}\hat{\Psi}_{elec} = E_{elec}\hat{\Psi}_{elec} \tag{2.5}$$

such that:

$$E_{tot} = E_{elec} + E_{nuc} \tag{2.6}$$

The wave function is not an observable and it cannot be measured. The only physical representation of the wave function is that the square of the wave function:

$$|\Psi_i(\vec{r}_1, ..., \vec{r}_N)|^2 d\vec{r}_1 ... d\vec{r}_N \tag{2.7}$$

gives the probability that N electrons are found simultaneously in the volume given by $d\vec{x_1}...d\vec{x_N}$. If we integrate over all space, the result must be unity since all the electrons must be found somewhere in space.

All electrons are fermions (spin=$\frac{1}{2}$). This means that the wave function must be anti-symmetric with respect to interchange of any two electrons. This applies to

both spatial and spin coordinates.

$$\Psi_i(\vec{r}_1, ..., \vec{r}_i, \vec{r}_j, ..., \vec{r}_N) = -\Psi_i(\vec{r}_1, ..., \vec{r}_j, \vec{r}_i, ..., \vec{r}_N) \tag{2.8}$$

### The Variational Principle

Except for a few simple cases, it is not possible to solve the Schrödinger Equation exactly. However, we can use the variational principle to systematically get closer to the ground state $\Psi_0$, and its corresponding ground state energy. The variational principle states that the expectation value of a Hamiltonian operator, $\hat{H}$, for any trial wave function $\Psi_{trial}$ will be higher than the energy of the ground state[1]:

$$\langle \Psi_{trial} | \hat{H} | \Psi_{trial} \rangle = E_{trial} \geq E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \tag{2.9}$$

where $E_{trial} = E_0$ only if $\Psi_{trial} = \Psi_0$.

The problem is then reduced to the search of the wave function that gives us the lowest corresponding energy. In theory, to reach the ground state all we need to know is the number of electrons, and the location and charge of all nuclei.

### The Electron Density

Keeping in mind the physical interpretation of the wave function, $\Psi$, we can define the electron density:

$$\rho(\vec{r}) = N \int ... \int |\Psi_i(\vec{x}_1, ..., \vec{x}_N)|^2 d\vec{s}_1 d\vec{x}_2 ... d\vec{x}_N \tag{2.10}$$

$\rho(\vec{r})$ determines the probability of finding, within a given volume, $N$ electrons, where one has an arbitrary spin, due to integrating over $d\vec{s}$, and the rest $N-1$ have an arbitrary position.

The electron density is an observable with a finite maximum, however, it has a 'cusp' like shape around the position of the nuclei. We can further extend this concept to define the pair density:

$$\rho_2(\vec{x}_1, \vec{x}_2) = N(N-1) \int ... \int |\Psi_i(\vec{x}_1, \vec{x}_2, ..., \vec{x}_N)|^2 d\vec{x}_3 ... d\vec{x}_N \tag{2.11}$$

this gives the probability of finding two electrons with spins $\sigma_1$ and $\sigma_2$ in the volume elements $d\vec{x_1}$ and $d\vec{x_2}$, with the rest of the electrons having arbitrary position and spin. Since electrons are fermions, the probability of finding two electrons with

[1] Using Dirac notation, where $\hat{O}$ is an operator and $\langle \hat{O} \rangle = \langle \Psi | \hat{O} | \Psi \rangle = \int ... \int \Psi \hat{O} \Psi^* d\vec{x}_1 ... d\vec{x}_N$

the same spin in the same position is zero. Hence, the position of one electron is correlated to the position of all the other electrons. We can define the 'exchange correlation hole', which is the difference between the conditional probability[2] and the uncorrelated probability:

$$h_{xc}(\vec{x}_1, \vec{x}_2) = \frac{\rho_2(\vec{x}_1, \vec{x}_2)}{\rho(\vec{x}_1)} - \rho(\vec{x}_2) = \rho(\vec{x}_2) f(\vec{x}_1, \vec{x}_2) \tag{2.12}$$

**The Hohenberg-Kohn Theorems**

**The First Hohenberg-Kohn Theorem:**

This first theorem (*40*) states that *the external potential $V_{ext}(\vec{r})$ is (to within a constant) a unique functional of $\rho(\vec{r})$; since, in turn $V_{ext}(\vec{r})$ fixes $\hat{H}$ we see that the full many particle ground state is a unique functional of $\rho(\vec{r})$.*

This implies that two different potentials, $V_{ext}(\vec{r})$ and $V'_{ext}(\vec{r})$, will never result in the same $\rho(\vec{r})$. Hence the complete ground state energy is a functional of the ground state electronic density.

$$E_0[\rho_0] = T[\rho_0] + E_{ee}[\rho_0] + E_{ne}[\rho_0] \tag{2.13}$$

We can divide the terms into components that are dependent on the specific system ($N$,$R_A$,$Z_A$) and ones that are not:

$$E_0[\rho_0] = \underbrace{T[\rho_0] + E_{ee}[\rho_0]}_{\text{independent HK functional}} + \underbrace{\int \rho_0(\vec{r}) V_{ne} d\vec{r}}_{\text{system dependent}} \tag{2.14}$$

**The Second Hohenberg-Kohn Theorem - Variational Principle:**

The second HK theorem is the equivalent of the variational principle described by Eqn. 2.9, but for the electron density. The theorem states that only the $\rho(\vec{r})$ that represents the true ground state of the system, $\rho_0(\vec{r})$, will deliver the ground state energy.

$$\langle \Psi_{trial} | \hat{H} | \Psi_{trial} \rangle = T[\rho_{trial}] + V_{ee}[\rho_{trial}] + \int \rho_{trial}(\vec{r}) V_{ext} d\vec{r} = E[\rho_{trial}] \geq \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \tag{2.15}$$

[2]The probability of finding an electron at $\vec{x_1}$ given an electron at $\vec{x_2}$

### The Kohn-Sham Approach

We saw that the energy of the ground state can be reached by minimising the universal HK functional and the system dependent component. However, the explicit form of the HK functional is the major challenge in DFT. The HK functional contains three contributions: kinetic energy, classical Coulomb interaction and the non-classical interaction:

$$F_{HK}[\rho] = T[\rho] + J[\rho] + V_{nc}[\rho] \tag{2.16}$$

The classical Coulomb interaction $J[\rho]$ is the only known component in $F_{HK}[\rho]$. The approach suggested by Kohn and Sham (*41*) is to calculate the kinetic energy of a reference, fictitious, system of non-interacting electrons with the same density $\rho(\vec{r})$ as the real system. The density is reformulated as a set of $N$ single-particle orthonormal functions, $|\psi_i\rangle$, and the kinetic energy can now by calculated exactly:

$$T_s = -\frac{1}{2}\sum_{i}^{N}\langle\psi_i|\nabla^2|\psi_i\rangle \tag{2.17}$$

the difference between this solution to the real kinetic energy is added to the difference between the classic potential and the non-classical one, introducing the exchange-correlation energy:

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (E_{ee}[\rho] - J[\rho]) \tag{2.18}$$

if the orbitals are chosen correctly, $E_{xc}[\rho]$ is usually small. This method computes exactly as much information as possible, leaving only $E_{xc}[\rho]$ as an approximate functional.

Minimising the Kohn-Sham functional results in the Kohn-Sham equations:

$$\Big(-\frac{1}{2}\nabla^2 + \Big[\int\frac{\rho(\vec{r'})}{|\vec{r}-\vec{r'}|} + V_{xc}(\vec{r}) - \sum_{A}^{M}\frac{Z_A}{|\vec{r}-\vec{R_A}|}\psi_i(\vec{r})\Big]\Big) = \varepsilon_i\psi_i(\vec{r}) \tag{2.19}$$

this set of equations is the key to uniquely determining the orbitals on our non-interacting system such that the reference system will have the exact same density as the real one.

If we know the form of $V_{xc}$ these equations will result in the exact ground state and energy. Unfortunately, this potential in unknown and must be approximated. We can therefore only define $V_{xc}$ as:

$$V_{xc} \equiv \frac{\delta E_{xc}}{\delta \rho} \tag{2.20}$$

**Exchange-Correlation Functionals**

All known exchange-correlation functionals are approximate. The success of DFT is dependent on the success of the chosen exchange-correlation functional. Unfortunately, there is no known 'one solution fits all' and the functional needs to be chosen according to the desired application.

Although the real form of the exchange-correlation functional is inherently non-local, a local potential is computationally favourable. The simplest exchange-correlation functional is the Local-Density Approximation (LDA), suggested by Kohn and Sham (*41*). LDA assumes that the exchange-correlation energy per electron, positioned at $r$ is the same as an homogeneous electron gas with the same density (per electron, at $r$):

$$E_{xc}[\rho(\vec{r})] = \int E_{xc}^{hom}[\rho(\vec{r})]\rho(\vec{r})d\vec{r} \tag{2.21}$$

Ceperley and Alder (*42*) took the suggested LDA form, and found an estimate for the exchange-correlation functional by solving the ground state of the uniform electron gas using the Monte Carlo method. This estimate was then parameterised for use in DFT methods (*43*).

The LDA approach is usually not accurate enough for many chemical applications. Typical problems with LDA includes over estimation of the binding energy and underestimation of the lattice parameters. The semi-local Generalised Gradient Approximation (GGA) includes a gradient correction to the electron density:

$$E_{xc}[\rho(\vec{r})] = \int E_{xc}^{hom}[\rho(\vec{r}), \nabla\rho(\vec{r})]\rho(\vec{r})d\vec{r} \tag{2.22}$$

This approximation is usually more accurate then LDA, but not always. Including the gradient can improve the total energy calculation since it reduces the error in bonding energy, but unlike LDA, GGA doesn't have one universal form. The Perdew-Bruke-Ernzerhof (PBE) functional (*44*) is one widely used parametrisation of GGA.

Traditional functionals do not describe the Van der Waals interactions. This may be acceptable since the van der Waals interactions tend to be significantly weaker then covalent or ionic interactions. However, for longer range interactions, such as the cohesion of layered materials or the creation of macromolecular assem-

blies in biological material, van der Waals interactions are a key factor. One way of resolving this issue is through Van der Waals functionals. These have been developed (*45–47*), but unfortunately have not yet produced chemically accurate results (*48*, *49*).

Another way to account for Van der Waals interactions is to use DFT-D schemes. DFT-D schemes correct the DFT energy with a semi-empirical dispersion correction that accounts for long range forces that bind the molecules together by accounting for the London dispersion forces. The form of the pairwise energy correction due to the London dispersion forces is:

$$\delta E_{AB}^{dis} = \frac{C_6}{R^6} \tag{2.23}$$

where $A$ and $B$ are a pair of atoms. The difference between the DFT-D schemes is in how the interatomic dispersion coefficients, $C_6$, are calculated.

The scheme used in this work was suggested by Tkatchenko and Scheffler (*50*) (denoted TS scheme). The TS scheme is a parameter-free method, where the $C_6$ coefficients are derived from the DFT electron density and reference data. This is to account for the effect of the local environment of the atoms on the values of the $C_6$ coefficients. Other examples for DFT-D schemes are the Grimme dispersion correction (G06 scheme) (*51*), OBS scheme (*52*) and the JCHS scheme (*53*).

Examples of DFT-D schemes applications, in particular TS and G06 schemes, are: study of the electronic structure and chemically bond of naphthalene and anthracene (*54*); side-stacking and electronic properties in thiophene–quinoxaline polymers (*55*); polymorphism in methyl paraben (*55*); and hydrogen bonding and $\pi$-$\pi$ interactions in indomethacin and nicotinamide co-crystals (*56*)

## 2.3 Solid State Nuclear Magnetic Resonance Theory

The material presented in this section is based upon the theory and derivations that are presented in references (*57–60*).

### 2.3.1 Magnetism

All substances interact with magnetic fields. Materials with a permanent magnetic moment are commonly known as magnets. The majority of substances, however, have induced magnetism, i.e., only possess a magnetic moment when in the presence of an external magnetic field.

The ratio between the induced magnetic moment, $\mu$, and the applied magnetic field $B_0$ is determined by $\chi$, the magnetic susceptibility of the material:

$$\mu \propto \chi B_0 \tag{2.24}$$

Materials with $\chi > 0$ are paramagnetic, and will tend to align with the magnetic field. Materials with $\chi < 0$ are diamagnetic, and attempt to reject the external field. Usually $|\chi_{para}| \gg |\chi_{dia}|$ for bulk materials.

Magnetism has three microscopic sources: (i) electromagnetic currents, (ii) intrinsic magnetic moment of electrons, (iii) intrinsic magnetic moment of the nucleus. The nuclear magnetic moment is directly related to the spin of the nucleus:

$$\mu = \gamma I \tag{2.25}$$

where $\gamma$ is the gyromagnetic ratio, an intrinsic property of every nuclear isotope. Most nuclei have a $\gamma > 0$ while some nuclei (and the electron) have $\gamma < 0$.

#### Spin Precession and the Larmor Frequency

The magnetic moment of a nucleus is aligned with the direction of the spin polarization ($\gamma > 0$) or opposite to it ($\gamma < 0$). The response of the spin polarisation to an external field is to precess around it - a cone shaped movement, keeping the angle between the spin and the magnetic field the same as it was when sample is first exposed to the field.

The frequency of the precession is the Larmor frequency:

$$\omega_0 = -\gamma B_0 \tag{2.26}$$

The precession will either be clockwise or counter-clockwise, depending on the sign of the gyromagnetic ratio, $\gamma$.

### 2.3.2 Quantum Mechanical Description

The interaction energy of a single nucleus in an external static magnetic field, if the nucleus has a non-zero spin, is given by the Zeeman Hamiltonian:

$$\hat{H} = -\hat{\mu} \cdot B_0 \tag{2.27}$$

where $B_0$ is the external magnetic field and $\hat{\mu}$ is the nuclear magnetic moment operator such that:

$$\hat{\mu} = \gamma \hat{I} \tag{2.28}$$

$\hat{I}$ is the spin angular momentum operator and $\gamma$ the gyromagnetic ratio, as previously mentioned, an intrinsic property of the nucleus. If we define the $\hat{z}$ axis in the direction of the magnetic field, we can write Eqn. 2.28 as:

$$\hat{H}_z = -\gamma \hat{I}_z \cdot B_0 = \omega_0 \hat{I}_z \tag{2.29}$$

where $\omega_0$ is the Larmor frequency (from Eqn. 2.26).

**Spin Angular Momentum Operator**

The spin angular momentum operator measures the spin angular momentum, when acted on a wavefunction. We can define four operators: $\hat{I}_x$, $\hat{I}_y$, $\hat{I}_z$ and $\hat{I}^2$, with the following relationships:

$$\hat{I}^2 = \hat{I}_x^2 + \hat{I}_y^2 + \hat{I}_z^2 \qquad [\hat{I}^2, \hat{I}_z] = 0 \qquad [\hat{I}_x, \hat{I}_y] = i\hbar \hat{I}_z \tag{2.30}$$

The eigenvalues of the $\hat{I}^2$ and $\hat{I}_z$ operators are (in $\hbar$ units):

$$\hat{I}^2|\Psi\rangle = l(l+1)|\Psi\rangle \qquad \hat{I}_z|\Psi\rangle = m|\Psi\rangle \tag{2.31}$$

where $m$ can range within the values $-l$,$-l+1$,...,$+l$. In the case of $l = \frac{1}{2}$, such as $^1$H and $^{13}$C nuclei, $m$ can be $-\frac{1}{2}$ or $\frac{1}{2}$. The Zeeman eigenstates and eigenvalues for a spin-half nucleus are:

$$\hat{H}_z|\uparrow\rangle = \frac{1}{2}\omega_0|\uparrow\rangle \qquad \hat{H}_z|\downarrow\rangle = -\frac{1}{2}\omega_0|\downarrow\rangle \tag{2.32}$$

commonly referred to as up ($|\uparrow\rangle$) and down ($|\downarrow\rangle$) states. A nucleus can also be in a state of super-position of the eigenstates:

$$|\Psi\rangle = C_\uparrow|\uparrow\rangle + C_\downarrow|\downarrow\rangle \tag{2.33}$$

$|C_\uparrow|^2$ and $|C_\downarrow|^2$ represent the probability of the spin to collapse to an up or down state when measured, respectively.

### 2.3.3 Density Operator

When describing an ensemble of nuclei, it is useful to define the density operator:

$$\hat{\rho} = \overline{|\Psi\rangle\langle\Psi|} \tag{2.34}$$

The over-bar notes an ensemble average of all the nuclear spin wavefunctions in the system (over-bar will be omitted from here onwards). For a system consisting of isolated half spin nuclei we can write $\hat{\rho}$ in a matrix representation:

$$\hat{\rho} = \begin{pmatrix} |C_\uparrow|^2 & \mathrm{C}_\uparrow C_\downarrow^* \\ \mathrm{C}_\downarrow C_\uparrow^* & |C_\downarrow|^2 \end{pmatrix} \tag{2.35}$$

The density-matrix must be Hermitian, with the diagonal elements being the expectation values of $\hat{I}_z$, which are always real and between $[0, 1]$. The off-diagonal elements are complex, and can be redefined as a real coefficient and a phase term:

$$C_{\downarrow\uparrow} = a_{\downarrow\uparrow} e^{i\phi_{\downarrow\uparrow}} \tag{2.36}$$

inserting back into the density matrix:

$$\hat{\rho} = \begin{pmatrix} |a_\uparrow|^2 & \mathrm{a}_\uparrow a_\downarrow e^{i(\phi_\uparrow - \phi_\downarrow)} \\ \mathrm{a}_\uparrow a_\downarrow e^{-i(\phi_\uparrow - \phi_\downarrow)} & |a_\downarrow|^2 \end{pmatrix} \tag{2.37}$$

The diagonal elements relate directly to the populations of the up and down states (the populations will be equal when not in the presence of a magnetic field). In a system of non-interacting spins with random phases, the off-diagonal terms average out to zero. The off-diagonal elements will only be non-zero in a system where a phase coherence exists between individual spins. This coherence can be created and detected in an NMR experiment.

### 2.3.4 Time Evolution of the Density Operator

To describe a spin system during an NMR experiment, one has to describe the evolution of the density operator in time. While the Schrödinger equation describes the time evolution of pure states, the Liouville-von Neumann equation determines the evolution in time of the density operator:

$$\frac{d\hat{\rho}(t)}{dt} = -i\hat{H}|\Psi\rangle\langle\Psi| + i|\Psi\rangle\langle\Psi|\hat{H} = -i[\hat{H}, \hat{\rho}(t)] \tag{2.38}$$

Under the approximation that the Hamiltonian is constant in time (mostly valid for short periods of time) the solution to Eqn. 2.38 becomes:

$$\hat{\rho}(t) = e^{-\hat{H}t}\hat{\rho}(0)e^{\hat{H}t} = \hat{U}(t)\hat{\rho}(0)\hat{U}^*(t) \tag{2.39}$$

$\hat{U}(t)$ is termed a 'propagator'.

### 2.3.5 Solid-State NMR Interaction Hamiltonians

The Hamiltonian required to describe a solid-state NMR experiment, $\hat{H}_{total}$, can be divided into six different components:

$$\hat{H}_{total} = \hat{H}_Z + \hat{H}_{rf} + \hat{H}_\sigma + \hat{H}_D + \hat{H}_J + \hat{H}_Q \tag{2.40}$$

$\hat{H}_Z$ and $\hat{H}_{rf}$ represent the external interactions, Zeeman and radio-frequency, respectively. The rest are internal interactions: magnetic shielding ($\hat{H}_\sigma$), The indirect magnetic interaction between the nuclear spins and the external magnetic field; dipolar coupling ($\hat{H}_D$), direct magnetic between the nuclear spins; J-coupling ($\hat{H}_J$), indirect interaction of the nuclear spins through the electrons; and quadrupolar interaction ($\hat{H}_Q$), electric interaction between nuclei with spin $> \frac{1}{2}$ and electric fields within the sample.

Each of these can be expressed as an independent operator. Each operator will have a special reference system - namely the principal axis system (PAS), where the operator will be diagonal, i.e., only the diagonal components are non-zero. Unfortunately, the different Hamiltonian components will often not have the same PAS.

Transferring between any two reference systems is achieved by a set of three rotations, defined by three Euler angles, $\alpha$, $\beta$ and $\gamma$:

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma) \tag{2.41}$$

To describe such rotations, it is convenient to express the interaction operators using spherical tensors instead of Cartesian tensors:

$$\hat{H} = \sum_{j=0}^{2}\sum_{m=-j}^{j}(-1)^m A_{jm}\hat{T}_{jm} \tag{2.42}$$

Only the coefficients, $A_{jm}$, change under rotation in space in the following way:

$$R(A_{jm}) = \sum_{m'=-j}^{j} (-1)^m D^j_{m,m'}(\alpha, \beta, \gamma) A_{jm'} \tag{2.43}$$

where $D^j_{m,m'}$ are known as Wigner rotation matrices.

### High Field Approximation

The high field approximation, also known as the secular approximation, relies on the externally applied magnetic field being large enough to make the Zeeman interaction the dominant term. The approximation treats all other interactions as perturbations of the Zeeman Hamiltonian. This allows contributions only from terms that commute with the Zeeman Hamiltonian. Specifically, only the $m = 0$ tensor components survive.

### Zeeman Interaction

The Zeeman Hamiltonian is:

$$\hat{H}_Z = \hat{I}\hat{Z}\hat{B}_0 = \omega_0 \hat{I}_z \tag{2.44}$$

As previously described in Eqn. 2.29. The direction of $\hat{Z}$ is the direction of the magnetic field. Under all temperatures achievable in NMR experiments (more then a fraction above absolute zero), the density operator at thermal equilibrium, corresponding to the start of an NMR experiment (assuming sufficient time for NMR relaxation to have occurred) will be:

$$\hat{\rho}_{eq} \propto \hat{I}_z \tag{2.45}$$

### Radio frequency

$B_0$ is not the only magnetic field applied in an NMR experiment (see Fig.2.1). An oscillating, weaker field, $B_1(t)$ is introduced by a radio frequency (rf) pulse:

$$B_1(t) = 2B_1 cos(\omega_{rf} t + \phi)\hat{x} \tag{2.46}$$

$B_1(t)$ is generated by an oscillating current passing through a coil wrapped around the sample. The oscillation frequency, $\omega_{rf}$, should be set equal to or very close to $\omega_0$.

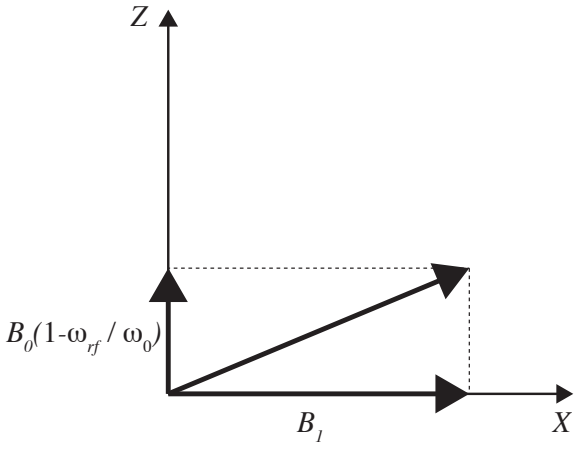


Figure 2.1: The magnetic fields present in the rotating frame of reference (at $\omega_{rf}$). Figure adapted from ref. (*61*).

To consider the effect of both $B_0$ and $B_1(t)$ on the density matrix, it is convenient to define a resonance offset frequency $\Omega = \omega_0 - \omega_{rf}$.

Since unlike $B_0$, $B_1(t)$ is time-dependent, the radio frequency component of the Hamiltonian is also time-dependent:

$$\hat{H}_{rf} = -\gamma B_1 \cos(\omega_{rf} t) = \omega_1 \hat{I}_x \tag{2.47}$$

with $\omega_1 = -\gamma B_1$ defined as the nutation frequency and for rf phase $\phi = 0$. In order to remove the time dependency of the Hamiltonian, a transformation into a rotating reference system is required, specifically, a reference system that is rotating about the $\hat{z}$ axis with a $\omega_{rf}$ frequency. $B_1(t)$ mixes the Zeeman states, but the eigenstates remain the Zeeman eigenstates. The time-dependent propagation of the density matrix for a system of isolated spin-half nuclei under an on-resonant pulse is:

$$\hat{\rho}(t) = e^{-\hat{H}_{rf}t}\hat{\rho}(0)e^{\hat{H}_{rf}t} = \frac{1}{2}\begin{pmatrix} \cos(\omega_1 t) & i\sin(\omega_1 t) \\ \text{-}i\sin(\omega_1 t) & \text{-}\cos(\omega_1 t) \end{pmatrix} \tag{2.48}$$

For a starting point of thermal equilibrium (see Eqn. 2.45) we can see the effect of different pulses: a $90°(\frac{\pi}{2})$ pulse makes diagonal terms (population) zero, and the off-diagonal (coherence) non-zero, while a $180°(\pi)$ pulse will cause the populations to invert.

With a starting point of transverse magnetism created by a 90° pulse, the density matrix will evolve in the following way:

$$\hat{\rho}(t) = \frac{1}{2}\begin{pmatrix} 0 & \mathrm{e}^{-i\Omega t} \\ \mathrm{e}^{i\Omega t} & 0 \end{pmatrix} \tag{2.49}$$

In an NMR experiment, the NMR signal corresponds to the induced current in a coil due to the precessing magnetisation. We can represent the recording of signal using quadruple detection (see section 2.4.1) to the operator $[\hat{I}_+ = \hat{I}_-^*]$, where $\hat{I}_+ = \hat{I}_x + i\hat{I}_y$ and $\hat{I}_- = \hat{I}_x - i\hat{I}_y$. Operating on the density operator will result in a real and imaginary part (neglecting any losses due to relaxation):

$$Tr(\hat{I}_+\hat{\rho}(t)) = \frac{1}{2}(\cos(\Omega t) + i\sin(\Omega t)) = \frac{1}{2}e^{i\Omega t} \tag{2.50}$$

**Magnetic Shielding**

In the presence of a magnetic field, the orbit of electrons in a diamagnetic sample will induce a magnetic field opposing the applied external field. The effective magnetic field in the sample will vary depending on the electronic, i.e. chemical, environment. The Hamiltonian corresponding to this interaction is:

$$\hat{H}_\sigma = \gamma \hat{I}\tilde{\sigma}B_0 \tag{2.51}$$

where $\tilde{\sigma}$ in the magnetic shielding tensor. In the spherical tensor representation (see Eqn.2.42), the $A_{00}^{PAS}$ component is rotationally invariant, and is proportional to the sum of the diagonal values of the shielding tensor. However due to the $A_{20}^{PAS}$ component, in an NMR spectrum of a powder (made of small crystals aligned in different orientations), we will get an anisotropic broadening of the spectral lines.

In an experiment, the measured quantity is the chemical shift, the relative offset of the measured frequency from a reference frequency, $\omega_{ref}$:

$$\delta_{iso} = \frac{\omega - \omega_{ref}}{\omega_{ref}} \times 10^6 \tag{2.52}$$

The chemical shift is a dimensionless parameter that is expressed in parts per millions (ppm). Note that this value is independent of $B_0$. $^1$H and $^{13}$C, are both referenced against the compound TMS (Tetramethylsilane), for example. Other nuclei are referenced against other standard materials.

**Dipolar Coupling**

Dipolar coupling is an interaction between nuclear spins that occurs through space. It decays as $\frac{1}{r^3}$, resulting in a noticeable effect only for a short distance, $r$. The

Hamiltonian of the dipolar interaction is:

$$\hat{H}_D = \hat{S}_1 \tilde{D} \hat{S}_2 \tag{2.53}$$

$\tilde{D}$ is a second-rank Cartesian tensor representing the interactions between $\hat{S}_1$ and $\hat{S}_2$, the interacting spins. We can define $b_{S_1S_2}$ (in units of $\frac{rad}{s}$) as:

$$b_{S_1S_2} = -\frac{\hbar\mu_0}{4\pi}\frac{\gamma_{S_1}\gamma_{S_2}}{r^3} \tag{2.54}$$

where $\mu_0$ is the permeability of space.

For a static experiment, in the spherical tensor representation, there is only a single component in the PAS:

$$\hat{H}_D^{PAS} = \sqrt{6} b_{S_1S_2} \hat{T}_{20} \tag{2.55}$$

and in the lab reference system for a static NMR experiment (see Eqn.2.43):

$$A_{20}^{LAB} = \sqrt{6} b_{S_1S_2} \frac{1}{2}(3\cos^2(\beta) - 1) \tag{2.56}$$

where $\beta$ is the second Euler angle when rotating between the reference systems as defined in Eqn. 2.41. In a powdered sample, $\beta$ will vary between a range of values due to the different crystallite orientations. As for the case of magnetic shielding, this will result in line broadening. In solution NMR these line broadening effects are eliminated by the rapid tumbling motion of the molecules, while in solid-state NMR, magic angle spinning (MAS) is used. Under MAS:

$$A_{20}^{MAS} = \sqrt{6} b_{S_1S_2} \frac{1}{2}(3\cos^2\theta_R - 1)(3\cos^2\beta - 1). \tag{2.57}$$

By choosing $\theta_R$ correctly, we can set $A_{20}^{MAS}$ to zero. MAS will be discussed further in section 2.4.4, i.e. such that $\cos\theta_R = \frac{1}{\sqrt{3}}$.

### 2.3.6 J Coupling

In addition to the though space dipolar interaction, nuclear spins have a through bond interaction: J coupling. When atoms are bonded, they share a pair of opposite spin electrons (due to the Pauli exclusion principle). The electrons are coupled to the nuclear spin through a hyperfine interaction creating two energy levels that manifest as a splitting in the spectral peaks. Although easily observed in solution-state NMR, J-coupling is usually on the order of 10-100 Hz, hence typically it will be lost in the line broadening the occurs in solid-state NMR. No direct measurement

of J-coupling was done in this work, however, the interaction is being utilized in experiments performed within this work, as further discussed in section 2.4.5.

### 2.3.7 Quadrupole Interaction

The Quadrupole interaction is the remaining internal interaction, however, it only affects nuclei with spin $> \frac{1}{2}$, hence does not affect $^{1}$H and $^{13}$C nuclei studied in this work and will not be discussed here further.

## 2.4 NMR experimental techniques and methods

### 2.4.1 1D NMR Experiment

A basic one dimensional experiment consists of applying a 90° ($\frac{\pi}{2}$) pulse which transfers the bulk magnetism the from the $\hat{z}$ axis (the direction of $B_0$) to the $x$ - $y$ plane (the direction of $B_1$). As mentioned in section 2.3.5, $\omega_{rf}$ needs to be equal or very close to $\omega_0$. The signal detected is then mixed with the spectrometer reference frequency, $\omega_{rf}$, such that the oscillation of the resonance offset frequency, $\Omega$, is observed.

#### Quadrature Detection

Signal detection is achieved using the quadrature detection technique - the signal is detected in the $x$ - $y$ plane, with both a real and an imaginary component of the NMR signal being recorded:

$$S(t) = (\cos(p\Omega t) - i\sin(p\Omega t))e^{\frac{-t}{T_2}} = e^{-ip\Omega t}e^{\frac{-t}{T_2}} \qquad t \geq 0 \tag{2.58}$$

where $p$ is the coherence order. $T_2$ is the time scale for the signal decay (so called transverse relaxation), and is typically in the order of milliseconds for organic solids. Note that for solids, $T_2$ is much smaller then $T_1$ (so called spin-lattice relaxation), the characteristic time the system required to return to thermal equilibrium. The effects of the signal decay are demonstrated in Fig. 2.2.

It is helpful to view the signal in the frequency domain:

$$S(\omega) = A(\omega) - iD(\omega) = \underbrace{\frac{T_2^{-1}}{T_2^{-2} + (\omega - \Omega)^2}}_{absorptive} - i\underbrace{\frac{(\omega - \Omega)}{T_2^{-2} + (\omega - \Omega)^2}}_{dispersive} \tag{2.59}$$

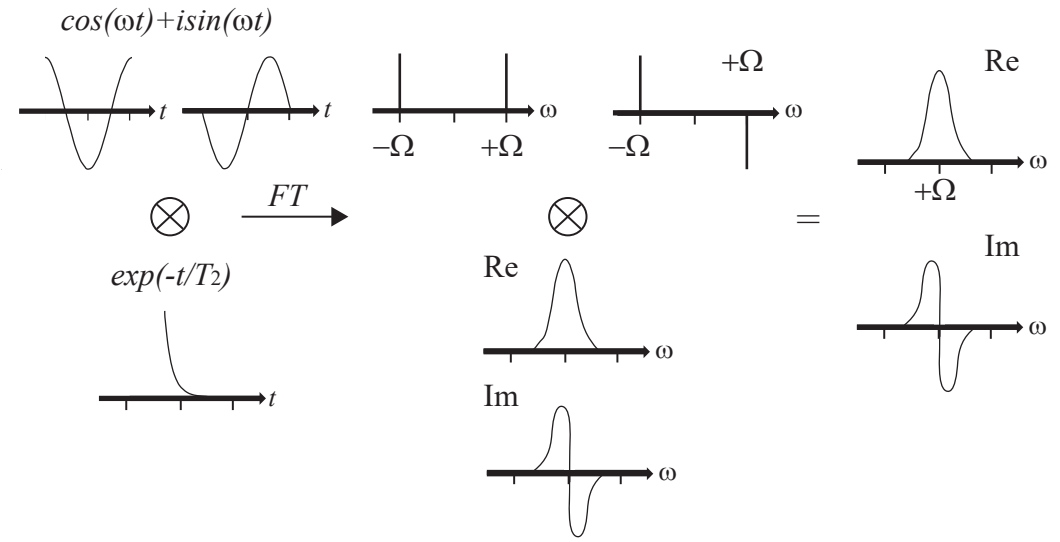


Figure 2.2: The effect of multiplying the signal $(\cos(\omega t) + i\sin(\omega t))$ by the decaying function $e^{-t/T_2}$ after a Fourier transform. The result is two line shapes, absorptive (Re) and dispersive (Im). Figure adapted from ref. (*61*).

The different lineshapes are illustrated in Fig. 2.2. Both $A(\omega)$ and $D(\omega)$ are centred at $\Omega$, however, $A(\omega)$ is maximal at $\omega = \Omega$ while $D(\omega = \Omega) = 0$. Since it is not practically possible to just disregard $D(\omega)$, the most desirable lineshape of the spectrum is achieved by adjusting the phase between the real and imaginary parts of the linear combination. This is referred to as 'phasing' the spectrum.

In practise the spectrometer uses a discrete Fourier Transform with the spectral width, $SW$, constrained by the inverse of the time between data points for acquisition (dwell time).

The only coherences directly detectable in a NMR experiment are single quantum coherences ($p = \pm 1$). Higher order coherences can be observed indirectly, which is discussed in the section below.

### 2.4.2 Multi-Dimensional NMR Experiments

The evolution in time of an ensemble average of a spin system is the observable in an NMR experiment. The system must be forced out of equilibrium in order to create a detectable change. To be specific, we can detect coherences between the eigenstates of the system as they evolve in time. However, as mentioned is the previous section, the evolution of non-observable coherences can be followed in the

indirect dimension(s), before the coherences with order of $\Delta p = \pm 1$ can be observed directly. In a two-dimensional (or higher) experiment, non-observable coherences are transferred into observable coherence for detection or heteronuclear correlation can be established. Fig. 2.3 presents a schematic pulse sequence illustrating this process: an initial rf pulse (preparation sequence) brings the system to an out-of-equilibrium state; afterwards, the system is allowed to evolve for a duration of $t_1$; the transfer pulse then transfers the $n$-order coherences into $-1$ coherences that are recorded during $t_2$.

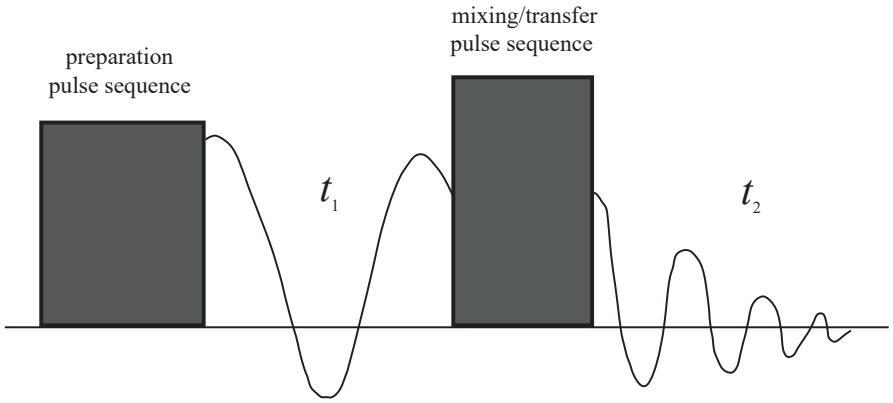


Figure 2.3: A schematic pulse sequence of a general 2D NMR experiment. Figure adapted from ref. (*61*).

The amplitude or the phase of the recorded signal is modulated according to the evolution during $t_1$. By repeating the experiment while varying $t_1$, a second dimension is generated. One dimension will contain information from the $-1$ coherences, and second from the $n$-order coherences. Selecting the desired order of coherence is achieved by phase cycling, as discussed in the section below.

The interpretation of a 2D spectrum is a correlation map. A peak at $(\omega_1, \omega_2)$ indicates that during $t_1$ there was an $n$-order coherence with a $\omega_1$ frequency, that transformed into a $-1$ coherence with a frequency of $\omega_2$.

This principle can be extended to higher-dimension experiments, but since these are not presented in this work, they will not be discussed further.

### Phase Cycling

The preparation sequence applied on a system of coupled spins can give rise to coherences with several orders. Phase cycling is used to cancel-out undesired coherences evolving in the system in addition to the desired coherence pathway.

The first golden rule of phase cycling is the following(*62*): *If the phase of a pulse is changed by $\phi$, a coherence undergoing a change in coherence level of $\Delta p$ acquires a phase shift of $\Delta p \cdot \Delta\phi$.* In a single phase cycle, the experiment is repeated $N$ times while incrementing the phase of the selected pulse(s) by $\frac{360^o}{N}$, and letting the receiver phase follow appropriately. Note that a receiver phase of $180^o$ means multiplying the signal by -1, while $90^o$ means switching the real and imaginary parts. The result is that a pathway that corresponds to a change in coherence order of $\Delta p \pm n \cdot N$,where $n$ is an integer, will prevail while others will be eliminated.

Table 2.2 demonstrates how a desired pathway $0 \rightarrow 2 \rightarrow +1 \rightarrow -1$, can be selected while an undesirable pathway, $0 \rightarrow -1 \rightarrow +1 \rightarrow -1$, is eliminated using phase cycling.

Table 2.2: Choosing a desired coherence pathway and eliminating an undesired one using phase cycling. Table adapted from ref. (*61*)

| cycle step | pulse phase | coherence phase | receiver phase | phase relative to receiver |
|---|---|---|---|---|
| desired pathway $0 \rightarrow 2 \rightarrow +1 \rightarrow -1$ | | | | |
| 1 | 0° | 0° | 0° | 0° |
| 2 | 90° | -180° | 180° | 0° |
| 3 | 180° | -360° | 0° | 0° |
| 4 | 270° | -540° | 180° | 0° |
| undesired pathway $0 \rightarrow -1 \rightarrow +1 \rightarrow -1$ | | | | |
| 1 | 0° | 0° | 0° | 0° |
| 2 | 90° | 90° | 180° | 90° |
| 3 | 180° | 180° | 0° | 180° |
| 4 | 270° | 270° | 180° | 270° |

### Quadrature Detection for Two-Dimensional Experiments

A flowchart describing the quadrature detection process for 2D experiments is presented in Fig. 2.4. In a 2D experiment, it is vital that the $\sin(\omega_1 t_1)$ and $\cos(\omega_1 t_1)$ components, corresponding to the first dimension, can be recorded separately, in addition to recording $\sin(\omega_2 t_2)$ and $\cos(\omega_2 t_2)$, corresponding to the second dimension, separately. If recorded together, the Fourier transform of the recorded signal will be:

$$e^{i\omega_1 t_1} e^{i\omega_2 t_2} \rightarrow (A_1 A_2 - D_1 D_2) + i(A_1 D_2 - D_1 A_2) \tag{2.60}$$

where $A_1 = A(\omega_1)$ and similarly for $A_2$, $D_1$ and $D_2$. In this case, the lineshape of the real part will be *phase-twisted*.

To achieve a purely absorptive lineshape in a two-dimensional experiment, two experiments are required. The preparation pulse sequence of the second ex-

periment is altered so that the coherence is phase shifted by 90° ($\frac{\pi}{2}$) between the experiments. This is achieved by shifting the phases of the preparation pulses by $\frac{\pi}{2}\Delta n$, where $\Delta n$ is the difference in coherence order from the start of the experiment to after the preparation pulse.

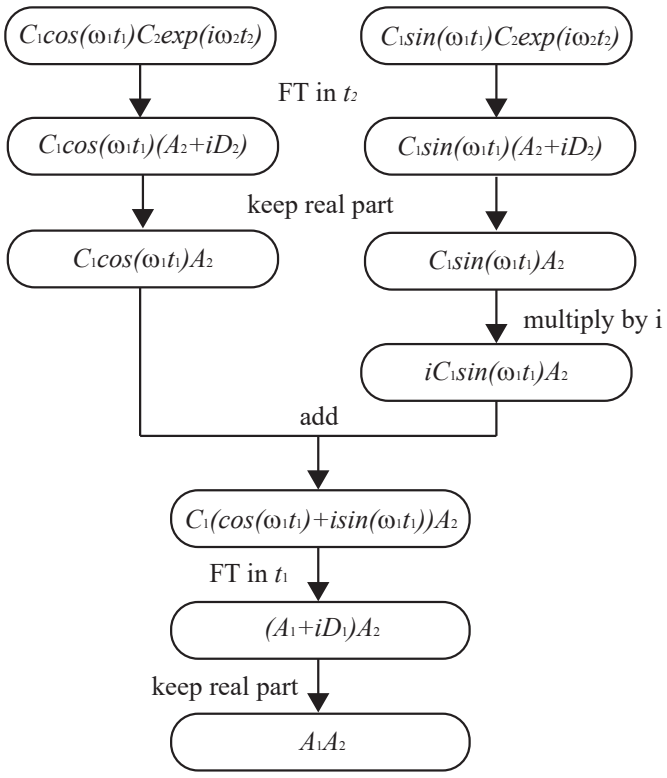


Figure 2.4: A flowchart describing the processing required to achieve a pure absorption lineshape from a 2D NMR experiment. Figure adapted from ref. (*61*).

### 2.4.3 The NMR spectrometer

An NMR spectrometer is an expensive, complex piece of equipment, designed to overcome the dual NMR challenges of sensitivity and resolution. An illustrative drawing of an NMR spectrometer with a sample inside in presented in Fig. 2.5. Fig. 2.6 is a schematic diagram of the instrumentation used in an NMR experiment.

Solid-state NMR experiments require a high, very homogeneous, magnetic field. This is to avoid line broadening due to slightly different applied magnetic field in different parts of the sample. In the majority of NMR spectrometers the magnet is a superconducting magnet. The advantage of superconductors is their ability, once charged with current, to keep running without an external power source. The

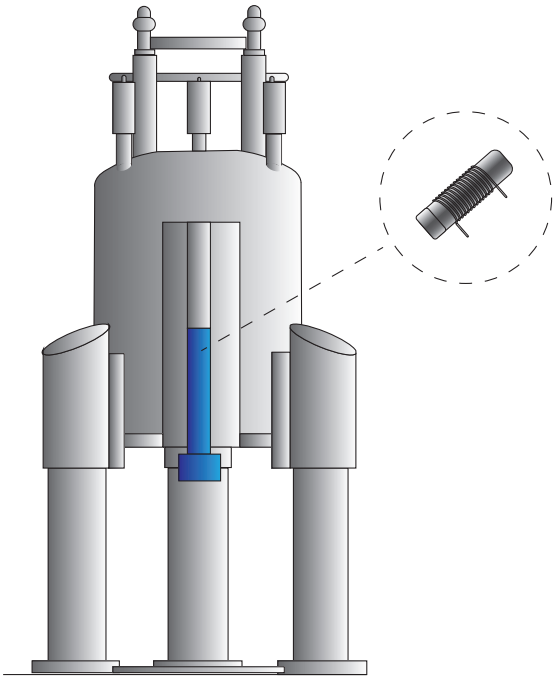

Figure 2.5: An illustrative drawing of an NMR magnet with a probe and a sample inside.

heart of the magnet, a Nb-Sn coil, is immersed in a bath of liquid He, surrounded in a liquid $N_2$ reservoir, to keep the magnet at the correct running temperature.

The sample itself is inserted into a probe, a separate device which is inserted into the spectrometer with the sample at the position of most homogeneous magnetic field. Inside the probe the sample is irradiated with rf waves, and the induced signal due to precession of magnetisation from the sample is detected. MAS probes spin the sample according to the desired limit and the capability of the probe. Probes usually also have temperature control features as well.

The signal detected in the coil inside the probe head is that of the carrier frequency plus the offset frequency, $\omega_0 + \Omega$. The signal is then amplified in the pre-amplifier. Afterwards, the signal is mixed down to oscillate around an intermediate frequency, $\omega_{IF} + \Omega$. Mixing the frequency down means that the spectrometer is required to deal with a smaller range of frequencies. This is done by mixing the signal with a spectrometer generated frequency of $\omega_0 - \omega_{IF}$ and a filter that only selects the component oscillating around $\omega_{IF}$.

The real and imaginary parts of the signal are then separated by routing two halves of the signal through different channels, 90° out of phase with each other.

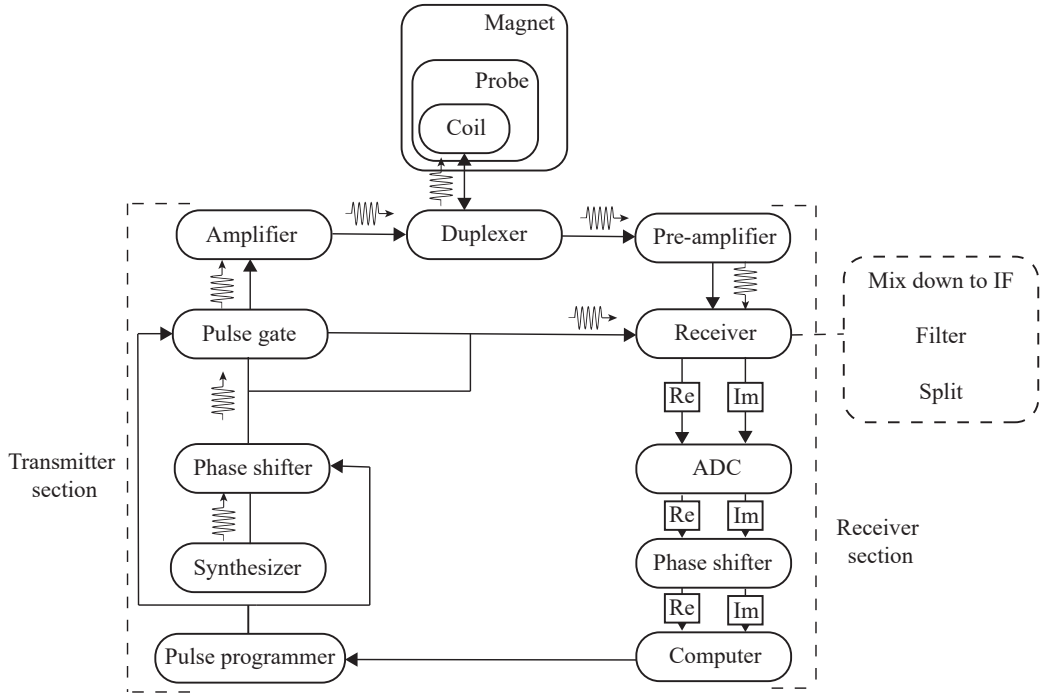


Figure 2.6: A schematic diagram of an NMR spectrometer, showing the console, magnet and probe. Adapted form ref. (*57*).

The complex signal is then converted to a digital signal in the 'analogue to digital converter' (ADC).

### 2.4.4 Magic Angle Spinning

In solid-state NMR experiments, the sample studied is usually a powdered sample, containing many crystallites in different orientations. Both magnetic shielding and dipolar interaction have a term with angular dependence of $(3\cos^2\beta - 1)$ (see Eqn. 2.56) where $\beta$ is the angle between the $\hat{z}$ axis of the shielding tensor PAS as defined in Eqn. 2.41 and the rotation axis (see Fig.2.7). As previously mentioned, in solution-state NMR these are averaged out due to the rapid tumbling motion of the molecules while in liquid state.

If the sample is spun at an angle of $\theta_R$ relative to the $\hat{z}$ axis in the lab frame it can be shown that, averaged over one rotor period:

$$\langle 3\cos^2\theta - 1\rangle = \frac{1}{2}(3\cos^2\theta_R - 1)(3\cos^2\beta - 1) \tag{2.61}$$

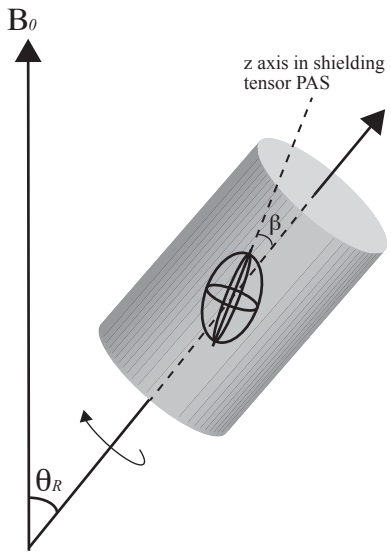


Figure 2.7: Rotation of a sample at the 'magic angle', $\theta_R$, during an MAS experiment. $\beta$ is the angle between the $\hat{z}$ axis of the shielding tensor PAS as defined in Eqn. 2.41.

If $\theta_R$ is set to 54.74°, the term in Eqn. 2.61 averages to zero.

#### Spinning side-bands

The spinning frequency should ideally be fast in comparison to the anisotropy of the interaction that we are trying to eliminate. The chemical shift anisotropy is proportional to the external applied field, $B_0$, therefore faster spinning is useful when a larger magnetic field is applied. If the MAS frequency is less then the linebroadening, 'spinning side-bands' are observed. These are sharp lines, separated by the spinning frequency, centred around the line of the isotropic chemical shift.

At the moment, commercial probes above 100 kHz are available, however, they come at the cost of a smaller rotor, hence less sample volume (for example, Bruker makes a 111 kHz probe for a 0.7 mm rotor).

### 2.4.5 Experiments Used in the Presented Work

#### Cross Polarization

Cross polarization (CP) is typically used to observe less abundant nuclei, such as $^{13}$C, by transferring magnetization from a more abundant nuclei, typically, $^{1}$H. The spin-half nucleus $^{13}$C makes only up 1.1% of the carbons when in natural abundance,

and combined with a lower gyromagnetic ratio, it is difficult to observe directly for two main reasons: first, low signal-to-noise ratio; second, long $T_1$ relaxation time due to lack of strong homo-nuclear dipolar interactions. $^1$H, on the other hand, is about 99.99% naturally abundant, giving a strong signal while suffering from lack of resolution due to strong homonuclear $^1$H-$^1$H dipolar coupling. Since the $^1$H relaxation time is usually much shorter then $^{13}$C, it is not required to wait until the $^{13}$C nuclear spins return to equilibrium between experiments, allowing a more rapid signal acquisition.

In the CP technique, i.e. application of rf irradiation on both nuclei simultaneously, the magnetisation is usually transferred from a nucleus with a higher gyromagnetic ratio (usually $^1$H), to a nucleus with a lower gyromagnetic ratio. The optimal enhancement gained is $\frac{\gamma_1}{\gamma_2}$ (4 for $^1$H and $^{13}$C). To achieve the transfer of magnetisation, the amplitude of the pulses must be set such that the Hartmann-Hahn condition is satisfied (*63*):

$$\gamma_1 B_1(S_1) = \gamma_2 B_1(S_2) \tag{2.62}$$

and with MAS:

$$\gamma_1 B_1(S_1) = \gamma_2 B_1(S_2) \pm n\nu_R \tag{2.63}$$

where $\nu_R$ is the MAS spinning frequency, and $\gamma_1$ and $\gamma_2$ are the gyromagnetic ratios of $S_1$ and $S_2$ nuclei, respectively. An efficient CP transfer is achieved by using a ramped pulse on the $^1$H channel such that the rf nutation frequency is increased gradually during the contact time pulses (*64*). See Fig. 2.8.

The efficiency of the transfer depends upon the strength of the dipolar interaction. This means that in a CP spectrum, the intensity of the peak is not purely an indication of the number of nuclear species, since it is weighted by a dependence on the corresponding carbon atoms' proximity to hydrogens.

### C-H correlation experiment - INEPT

In the INEPT (insensitive nuclei enhanced by polarization transfer) (*65*, *66*) experiment, magnetization is transferred between nuclei of bonded atoms. Similar to the one-dimensional CP experiment, the maximal gain in sensitivity is the ratio of gyromagnetic ratios: $\frac{\gamma_1}{\gamma_2}$, which is 4 for $^1$H and $^{13}$C. The difference is that in an INEPT experiment, a J coupling is used to transfer the magnetization rather than a dipolar coupling.

The 2D refocused INEPT experiment has been adapted for solid-state NMR

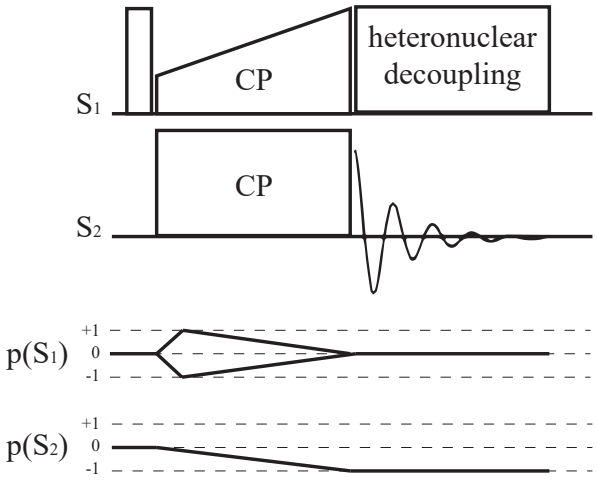


Figure 2.8: Pulse sequence and coherence transfer pathway diagram for a cross polarization (CP) experiment. The magnetization is transferred from the high $\gamma$ nuclei, $S_1$ to the lower $\gamma$ nuclei, $S_2$.

by adding homonuclear decoupling that, in addition to MAS, reduces the $^1$H dipolar interaction (*67*). Only bonded $^{13}$C - $^1$H pairs will appear in the spectrum, and when used with a short $\tau$ period, the experiment is selective only to pairs linked by a single bond. A refocused INEPT pulse sequence and coherence diagram is presented in Fig. 2.9. Transverse magnetization is allowed to evolve during $t_1$. As seen in Fig. 2.9, the heteronuclear transfer occurs via the second $\frac{\pi}{2}$ pulse on the $^1$H channel. The second heteronuclear spin-echo block ($\tau'$ - $\pi$ - $\tau'$) refocuses the signal which results in in-phase peaks, that do not cancel each other out. In solid-state NMR experiments, $\tau$ and $\tau'$ are optimized for each sample.

### DUMBO Homonuclear Decoupling Scheme

DUMBO (decoupling using mind boggling optimization) (*68*, *69*) scheme uses on-resonance rf pulses to apply homonuclear decoupling. A series of pulses with the same frequency are applied with a varying phase. This is based on the BLEW-12 scheme (*70*). eDUMBO-$1_{22}$ (*71*) is an optimized version of fast MAS.

### $^1$H - $^1$H Double Quantum Experiment

$^1$H double quantum (DQ) correlation experiments are used to identify inter-molecular proximities between $^1$H pairs. The experiment probes both inter-molecular and intra-molecular distances, and it is not possible to distinguish between one and the

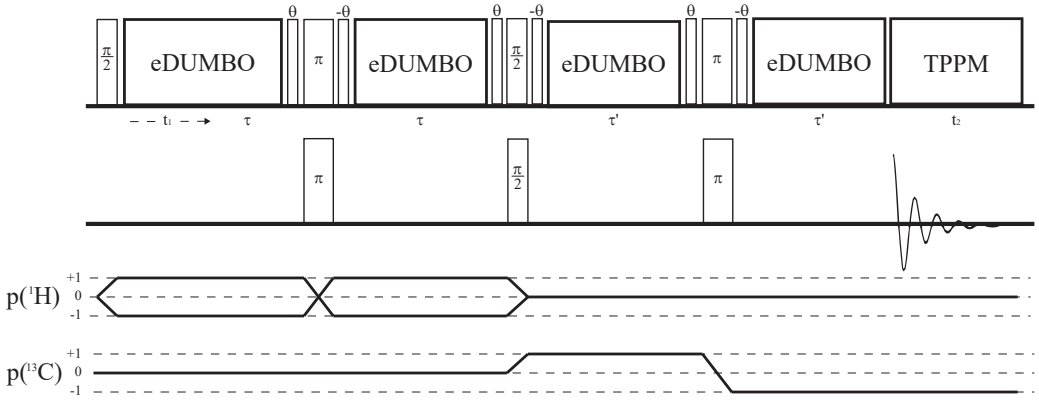


Figure 2.9: Pulse sequence and coherence transfer pathway diagrams for a 2D $^{13}$C - $^{1}$H refocussed INEPT experiment using eDUMBO-$1_{22}$ homonuclear decoupling. Figure is adapted from ref. (*67*).

other.

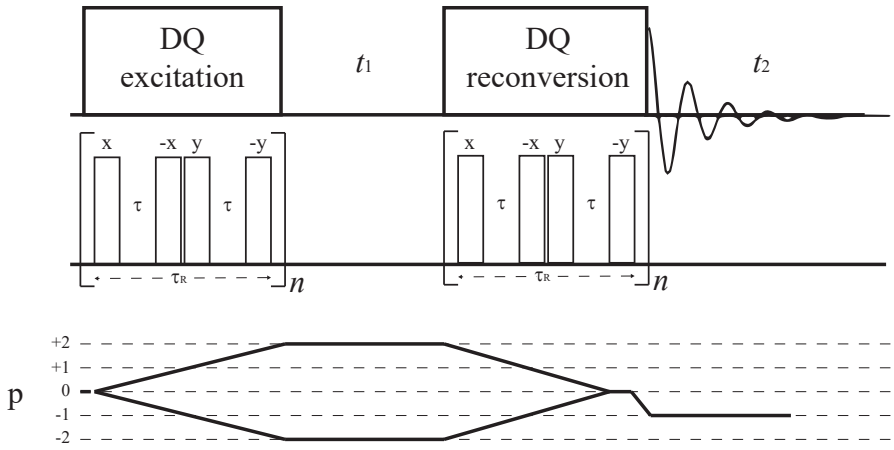


Figure 2.10: Pulse sequence and coherence transfer pathway diagram for a $^{1}$H - $^{1}$H DQ experiment using BABA dipolar recoupling.

All the $^{1}$H - $^{1}$H DQ experiments presented in this work use the Back-to-Back (BABA) dipolar recoupling scheme (*72*, *73*). The scheme is used to excite and reconvert DQ coherence ($\Delta p = \pm 2$). The basic BaBa scheme consists of four 90° pulses and two periods of free evolution per rotor period ($\tau_R$). Repeating the BaBa cycle over several rotor periods corresponds to longer dipolar recoupling times. A schematic pulse sequence and coherence transfer pathway diagram is presented in Fig. 2.10 with $n$ noting the number of cycle repetitions.

It was demonstrated that under MAS, due to time reversal symmetry, a simple 90 - $\tau$ - 90 pulse sequence will result in a disruptive interference that eliminates the signal after a single rotor period ($\tau_R$). In the BaBa sequence, a 90° phase shift of the two pulses on the second half of the sequence is introduced to overcome this obstacle (*74–76*).

The result of the experiment is a two-dimensional spectrum, where peaks only appear for coupled pairs. Usually, only peaks for pairs of $^1$H nuclei closer then 3.5 Å will appear in the spectrum (*74*). A schematic diagram of the spectrum is presented in Fig. 2.11.

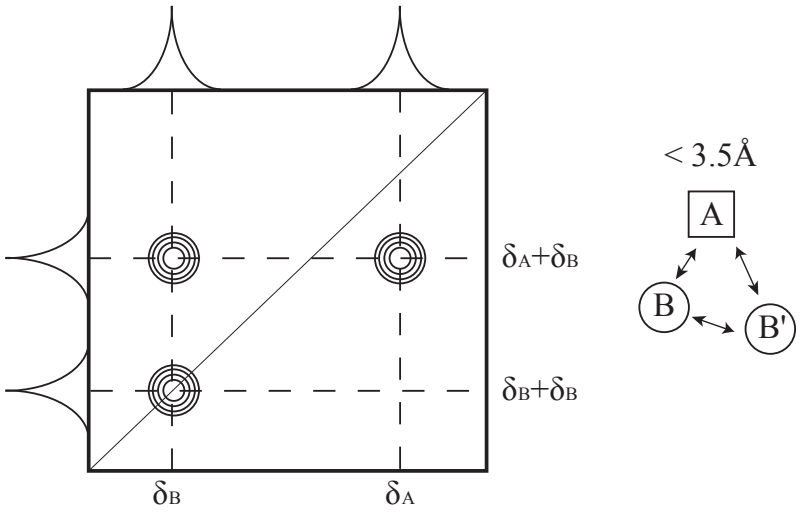


Figure 2.11: A schematic spectrum for a $^1$H - $^1$H DQ experiment, showing a correlation map for 2 nuclei that are separated by less than 3.5 Å. In this simple case, $A$ and $B$ represent different nuclei, and $B'$ is a symmetry copy of $B$. The distances between $A$ and $B$, $A$ and $B'$ and $B$ and $B'$ are all under 3.5 Å. In the SQ dimension, the peaks appear at the isotropic chemical shift, while in the DQ dimension, they are at the sum of the chemical shifts.

If homonuclear decoupling is used, the chemical shielding is also scaled in both the $^1$H SQ and DQ dimensions of the spectrum. In order to calibrate the chemical shift axes correctly, an additional 1D spectrum recorded using fast MAS alone is required, and is used to re-scale the spectrum.

## 2.5 Computation of Magnetic Resonance Parameters

### 2.5.1 GIPAW and CASTEP

#### Infinite Periodic Crystals

While macroscopic crystals are not actually infinitely large, they do contain a very large number of repeating units, the unit cell, such that most atoms within the crystal effectively 'feel' the environment of an infinite crystal.

Performing a simulation on all of the electrons in a real sample is impossible, but luckily, also unnecessary due to the periodic nature of the crystal. We can exploit the transitional symmetry that exists in all crystalline materials and focus a calculation on a single unit cell. By applying periodic boundary conditions we are in fact calculating the values for an infinite periodic crystal.

#### Bloch's Theorem

Calculating an infinite crystal from a single unit cell is possible because of Bloch's theorem: If the potential $V(\vec{r})$ is periodic on a lattice such that $V(\vec{r}) = V(\vec{r} + \vec{a_i})$ where $a_i$ is a lattice vector, then we can write the eigenstate of the single-particle Hamiltonian as:

$$\psi_{\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{\vec{k}}(\vec{r}) \tag{2.64}$$

where $u_{\vec{k}}(\vec{r})$ has the same periodicity as $V(\vec{r})$: $u_{\vec{k}}(\vec{r}) = u_{\vec{k}}(\vec{r} + \vec{a_i})$ for all lattice vectors $a_i$.

These are called 'Bloch states' and are labelled by their crystal momentum $\vec{k}$. Unique values of $\vec{k}$ only exist within one unit cell of the reciprocal lattice, or within the 1st Brillouin Zone (also known as the Wigner-Seitz cell in the reciprocal lattice). Fig. 2.12 gives a schematic representation of the unit cell, the reciprocal cell and the k-point sampling grid.

#### Planewave Basis Set

When attempting to solve the Kohn-Sham equations Eqn. (2.19) numerically, we are presented with the choice of a basis set for the Khon-Sham eigenstates. Choosing a planewave basis set, we can express the eigenstates as follows:

$$\psi_{\vec{k}}(\vec{r}) = \sum_i c_{\vec{k}}(\vec{G_i}) e^{i(\vec{k}+G_i)\cdot\vec{r}} \tag{2.65}$$

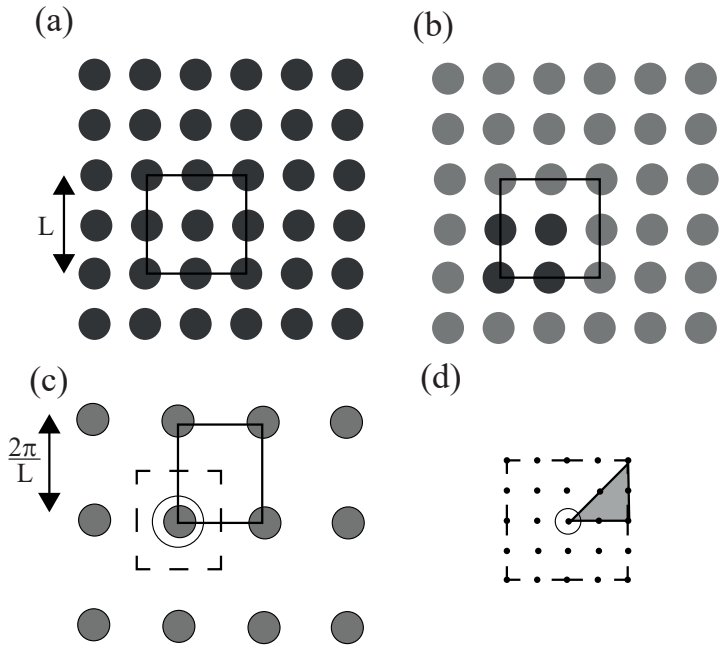


Figure 2.12: Figure adapted form Ref. (77). (a) The unit cell in real space, with the lattice shown in black. (b) The darker atoms are the unique atoms that must be included in the calculation. (c) The reciprocal cell, with the gamma point (the origin) indicated by a circle and the first Brillouin zone marked by a dashed line (d) The first Brillouin zone divided by a uniform grid of sampling points, though because of symmetry considerations, only the grid points in the triangle need to be calculated explicitly.

where $G_i$ are the reciprocal lattice vectors. $c_{\vec{k}}(\vec{G}_i)$ are determined by solving Eqn. (2.19).

There are many possible choices for a basis set, however, planewaves are naturally periodic and converge systematically as a function of the size of the basis, making them very suitable as a basis set for the Kohn-Sham eigenstates in crystalline solids.

Since the solution is numerical, we must limit the size of the basis set. This is achieved by only summing over $\vec{G}_i$ contained within a sphere with a given radius such that:

$$\frac{\hbar^2|\vec{k}+\vec{G}|^2}{2m} \leq E_{cut} \tag{2.66}$$

The larger the cut-off energy, $E_{cut}$, is, the more accurate the solution will be, but with the cost of the calculation being more computationally heavy. Transferring a

function between reciprocal and real space is achieved via Fourier Transform. Numerically efficient Fast Fourier Transform (FFT) is very important to the successful application of the formalism since it allows the calculating of different components in their most efficient basis, for example, the kinetic component in the reciprocal space, and the local part of the potential in real space.

### The Pseudopotential Approximation

In order to improve the computational efficiency of the calculations, the electronic wave functions should be represented by the smallest feasible basis set. This will minimise the cost of computational operations, FFT for example, and reduce the required memory capacity. The size of the basis set depends on the cut-off energy, $E_{cut}$, such that smoother wave function will require a lower $E_{cut}$ to be described accurately.

We can divide the electron in the atoms into two groups: core electrons, tightly bound to the nucleus, and the valence electrons, further away from the nucleus. An alternative, more practical definition is that all interactions between atoms are dictated by the valence electrons, while the core electrons have little to no effect outside the immediate environment of the nucleus. The calculated chemical properties depend on the overlap of the valance electron's wave functions, while the core electrons, to a good approximation, remain unchanged from their atomic gas nature when in the solid state.

If we were to run a calculation including all of the electrons the requirement that the electronic states must be orthogonal means that strong oscillations will occur for wavefunctions in the vicinity of the nucleus. These oscillations require a large basis set to describe properly, making the calculation computationally demanding while having little to no effect on the values of the properties we are trying to calculate.

In the pseudopotential approximation we only consider the valence electrons explicitly. The nuclear Coulomb potential is replaced by a smother pseudopotential and the core electrons are subsequently removed. The pseudopotential approximation thus improves the efficiency of the calculation in two ways: by reducing the number of electrons accounted for, and by reducing the number of basis states required to describe the valence electrons since the strong oscillations number have now been eliminated. Many types of pseudopotential have been proposed, however, the current state-of-the-art are ultrasoft pseudopotentials (USP) introduced by Vanderbilt in 1990 (*78*). USPs typically give converged results with a smaller basis set than other types of pseudopotential. A recent Science paper (*79*) examined

the reproducibility of DFT calculations of solids. It compared the performance of a wide variety of DFT codes and concluded that well constructed pseudopotentials give almost identical results to the best all-electron methods.

**Biot-Savart Law and Current Density**

The following sections mainly follow Chapter 3 of Ref. (*80*). Under the influence of a uniform external magnetic field, induced currents flow throughout a material. In a non-magnetic insulator, only the electrons contribute to the currents, via their orbital motion. We can define the current density $j(r)$, and the magnetic field it induces inside the material:

$$B_{in}(r) = \frac{\mu_0}{4\pi} \int d^3r' j(r') \times \frac{r - r'}{|r - r'|^3} \tag{2.67}$$

This is the Biot-Savart Law (in SI units). The magnetic shielding tensor, previously mentioned in Eqn. 2.51, gives the ratio between the induced and the external fields:

$$B_{in} = -\tilde{\sigma} B_{ext} \tag{2.68}$$

The key to calculating the shielding tensor is finding the current density. Since the induced field is small compared to the Zeeman interaction, the current can be calculated using perturbation theory. Keeping only terms linear in the external field:

$$j^{(1)}(r') = 2\sum_i [\langle \Psi_i^{(0)} | \hat{J}_{para}(r') | \Psi_i^{(1)} \rangle] + [\langle \Psi_i^{(1)} | \hat{J}_{para}(r') | \Psi_i^{(0)} \rangle] + 2\sum_i [\langle \Psi_i^{(0)} | \hat{J}_{dia}(r') | \Psi_i^{(0)} \rangle] \tag{2.69}$$

the $i$ subscript notes a summation over all occupied orbitals, the superscript notes the perturbation order (i.e. $j^{(1)}(r')$ is the induced current to first order in the applied field) and $\hat{J}(r)$ is the current operator, divided into paramagnetic and diamagnetic components:

$$\hat{J}(r') = \hat{J}_{dia}(r') + \hat{J}_{para}(r') \tag{2.70}$$

where

$$\hat{J}_{dia}(r') = \frac{e^2}{m_e} A(r') |r'\rangle\langle r'| \qquad \hat{J}_{para}(r') = -\frac{e^2}{2m_e} p|r'\rangle\langle r'| + |r'\rangle\langle r'|p \tag{2.71}$$

Using a symmetric gauge, $A(r) = \frac{1}{2} B \times r$, the derivation of the current

density results in:

$$j^{(1)}(r') = \frac{4e}{m_e}\sum_i Re[\langle\Psi_i^{(0)}|\hat{J}_{para}(r')\mathcal{G}(\varepsilon_i^0)r\times p|\Psi_i^{(1)}\rangle]\cdot B - \frac{e^2}{2m_e}\rho(r')B\times r' \quad (2.72)$$

where $\rho(r')$ is the ground state charge density and $\mathcal{G}(\varepsilon)$ is the Green function:

$$\mathcal{G}(\varepsilon) = \sum_e \frac{|\Psi_e^{(0)}\rangle\langle\Psi_e^{(0)}|}{\varepsilon - \varepsilon_e}. \quad (2.73)$$

**Calculating Magnetic Shielding in a Periodic System**

In the derivation above, there is no specific mention of the size of the system, and it is, in principle, possible to calculate the current density for a finite system. With a careful look at Eqn. 2.72, it is clear the second term will diverge with system size, due to the presence of the position operator. This seemingly nonphysical divergence is countered by the divergence of the first term, making only the sum of the two terms convergent and well defined. It is possible to rewrite Eqn. 2.72 as:

$$j^{(1)}(r') = \frac{4e}{m_e}\sum_i Re[\langle\Psi_i^{(0)}|\hat{J}_{para}(r')\mathcal{G}(\varepsilon_i^0)(r-r')\times p|\Psi_i^{(1)}\rangle]\cdot B - \frac{e^2}{2m_e}\rho(r')B\times r'. \quad (2.74)$$

Due to the local nature of the Green function, the current density remains finite.

When attempting to compute the current density for a periodic system, it is desirable to use the Bloch states (Eqn. 2.64). The difficulty is that the position operator does not have the periodicity of the unit cell. A solution to this problem was presented by Mauri et al. (*81*), who showed that it is possible to consider the response to a magnetic field with a finite wavelength, $q$, with the uniform field response recovered in the limit that $q \to 0$. In practice, this enables the use of the Bloch states, at the cost of an additional six calculations for each calculated point in the Brillouin zone ($k \pm q$ in each axis). The full derivation appears in Refs. (*7*, *8*, *81*).

The calculation of the induced field and the density current is performed in the reciprocal space such that:

$$B^{(1)}(G) = \mu_0\frac{iG\times j^{(1)}(G)}{G^2} \qquad B^{(1)}(R) = \sum_G e^{iG\cdot R}B^{(1)}(G) \quad (2.75)$$

The majority of the computational cost goes into calculating the current density, once that is achieved, calculating the shielding at multiple points is rapid. This is

useful when calculating the independent chemical shift (NICS) - the shielding in an arbitrary point on space, i.e, not in a centre of a nuclei.

When attempting to perform a calculation on an infinite periodic crystal, the $B^{(1)}(G=0)$ cannot be calculated. Indeed, this is not a bulk property (*81*, *82*), but an effect caused by surface currents, and will depend on the shape of the sample. Chemicals shifts found through MAS experiments, correspond to a spherical sample, regardless of the physical shape of the sample. As such, when calculated, the $G=0$ element is approximated as(*82*):

$$B^{(1)}(G) = \frac{2\mu_0}{3}\chi B \tag{2.76}$$

where $\chi$ is the macroscopic magnetic susceptibility. This can be calculated from first-principles following Ref. (*82*).

### Core Electrons Contributions and the Projector Augmented Wave Method

By reducing the computational cost of the calculations, the use of pseudopotentials makes calculations on larger systems feasible. The pseudopotential approximation is valid for the calculation of total energies, bond lengths and unit cell parameters. However, the wave functions of the valence electrons are not physical in the region near the nuclei. This results in errors when calculating proprieties that are sensitive to the shape of the wave-function near the nucleus, such as the magnetic shielding. The solution, introduced by Van de Walle and Blöchl is the Projector Augmented Wave (PAW) approach (*83*). In the PAW method, the all-electron wavefunction $\Psi$ is constructed from the pseudo wave function $\tilde{\Psi}$ by a linear transformation, $\mathcal{T}$

$$|\Psi\rangle = \mathcal{T}|\tilde{\Psi}\rangle \tag{2.77}$$

where

$$\mathcal{T} = \mathbf{1} + \sum_{\mathbf{R},n}[|\phi_{\mathbf{R},n}\rangle - |\tilde{\phi}_{\mathbf{R},n}\rangle]\langle\tilde{p}_{\mathbf{R},n}| \tag{2.78}$$

$|\phi_{\mathbf{R}n}\rangle, |\tilde{\phi}_{\mathbf{R}n}\rangle$ are all-electron and pseudo partial waves and $\langle\tilde{p}_{\mathbf{R},n}|$ are a set of projectors such that $\langle\tilde{p}_{\mathbf{R}n}|\tilde{\phi}_{\mathbf{R}'m}\rangle = \delta_{\mathbf{R}\mathbf{R}'}\delta_{nm}$. $\mathbf{R}$ is the atomic centre that the projector or partial wave is centred on, and $n$ is an index which refers to both the angular momentum quantum numbers and to an additional number which is used to differentiate between two projectors in the same angular momentum channel. In summary, the PAW method replaces the pseudo-component of the wavefunction for its all-electron form for each atomic state.

$\Omega_{\mathbf{R}}$ is the augmentation region, defined separately for each atom such that

the following conditions are met: a) outside $\Omega_{\mathbf{R}}$ the pseudo and all-electron partial waves are the same and the projectors $\langle \tilde{p}_{\mathbf{R}n}|$ are zero, b) within $\Omega_{\mathbf{R}}$, the $|\phi_{\mathbf{R}n}\rangle$ form a complete set for the valence wavefunctions. The augmentation region has no intrinsic physical significance. It must be small enough that the augmentation regions of different atoms do not overlap. However, as the augmentation region is made smaller, higher energy planewaves will need to be added to the basis to give converged results. In the PAW approach, every all-electron operator, $O$, is transformed to a corresponding pseudo operator $\tilde{O}$:

$$\tilde{O} = O + \sum_{\mathbf{R},n,m} |\tilde{p}_{\mathbf{R},n}\rangle[\langle\phi_{\mathbf{R},n}|O|\phi_{\mathbf{R},m}\rangle\langle\tilde{\phi}_{\mathbf{R},n}|O|\tilde{\phi}_{\mathbf{R},m}\rangle]\langle\tilde{p}_{\mathbf{R},m}| \tag{2.79}$$

The result is that the pseudo-operator, operating on pseudo wavefunctions, will give the same matrix elements as the all electron operator, acting on its corresponding wave functions.

### Gauge Including Projector Augmented Wave Method

Unfortunately, the PAW method is not practical in the presence of an external magnetic field, since an unrealistically large number of projectors are required to make the shielding transitionally invariant. This was rectified by the Gauge Including Projector Augmented Wave Method (GIPAW) presented by Pickard and Mauri (*7*). The GIPAW method offers a generalised version of the PAW transformation operator (Eqn.2.78) which imposes the transitional invariance explicitly:

$$\mathcal{T} = \mathbf{1} + \sum_{\mathbf{R},n} e^{\frac{ie}{2\hbar}r\cdot R\times B}[|\phi_{\mathbf{R},n}\rangle - |\tilde{\phi}_{\mathbf{R},n}\rangle]\langle\tilde{p}_{\mathbf{R},n}|e^{\frac{ie}{2\hbar}r\cdot R\times B}. \tag{2.80}$$

A computationally efficient method for combining the GIPAW method with ultrasoft pseudopotentials was presented at Yates et al. (*8*).

The starting point of a GIPAW calculation is the crystal structure. The initial unit cell is usually derived from experimental data via XRD (single crystal of powder) or from first principles through crystal structure prediction (CSP) methods. It is common practice to perform a geometry optimisation on the structure first. This will move the position of the atoms until the structure reaches a local minima. This stage is important for the positions of the hydrogen atoms in particular. Geometry optimisation can be done while allowing the unit cell dimensions to vary, or keeping them fixed. After the desired geometry has been found, the NMR tensors are calculated using the GIPAW method. The GIPAW method has been implemented into the CASTEP code (*84*), used in this work, and is also

available within the QUANTUM ESPRESSO(*85*) code. GIPAW has been applied extensively in order to study both organic and inorganic structures, with the studied nuclei typically being $^{1}$H, $^{13}$C, $^{15}$N and $^{17}$O for organic compounds. An extensive review of GIPAW applications can be found in Bonhomme *et al.* (*77*). A particular mention is given here to studies of the effect of hydrogen bonds and $\pi$-$\pi$ interactions on the $^{1}$H chemical shifts (*56*, *86–88*) as well as GIPAW studies of polymorphism in pharmaceutical drugs (*89–93*).

### Accuracy and Efficiency of GIPAW Calculations

When performing a first-principles calculation, the calculation parameters should be chosen in such a way that maximal efficiency is achieved without compromising the accuracy of the calculation for the desired application. We can distinguish between two types of approximations that affect the accuracy of the calculation: physical approximations include the exchange-correlation functional, neglect of thermal motion, frozen nuclei and core electrons, etc; and computational approximations, such as the size of basis set, and the choice of $k$-points. The effect of physical approximations can be evaluated by comparing the results to experimental data, while computational approximations can be systematically improved.

Computational approximations are tested through their impact on the convergence of the calculated parameters. The size of the basis set, for example, is determined by the energy cut-off, $E_{cut}$, as seen in Eqn. 2.66. The accuracy of the results will improve with the size of $E_{cut}$, until convergence is reached, i.e. the values do not change when $E_{cut}$ is increased further. It is worth noting that, usually, NMR shielding calculations require a higher energy cut-off to converge than a geometry optimisation. Similarly, the accuracy of the calculation will depend on the spacing between the $k$-points used for sampling the Brillouin zone. A Monkhorst-Pack grid (*94*) is an methodological way of choosing a set of $k$-points, with $0.1 \times 2\pi\text{Å}^{-1}$ being a sensible choice for spacing between the $k$-points. Both the value of $E_{cut}$ and the set of $k$-points should be chosen through a convergence test.

The accuracy of the results is only as good as the choice of the exchange-correlation functional. In this work the GGA functional PBE (*44*) is used with the TS (*50*) dispersion correction scheme employed for geometry optimisations. PBE usually produces good agreement with experiments and is used in the majority of GIPAW calculations. Chemical shifts will usually be with errors of 2-3%, though some exceptions have been found for $^{19}$F (*95*, *96*) and $^{17}$O (*97*). Two other important physical approximations are relativity and thermal motion, both of which have been neglected in this work.

## 2.6 Crystal structure prediction

The aim of crystal structure prediction (CSP) is to find the possible polymorphs, co-crystals, salts, hydrates etc. for a given compound with no experimental data except for the chemical formula of the molecule. There can be many potential applications for a successful CSP method, but there is a clear link to the pharmaceutical industry, with the hope that CSP methods can aid the expensive and long solid form screening process (*98*).

The progress made in the field of CSP (*99–101*) had been monitored by the 'blind tests of organic crystal structure prediction methods', run by the Cambridge Crystallographic Data Centre (CCDC) since 1999. The results of most recent blind test, the $6^{th}$ blind test, were published in 2016 (*102*). All the blind tests run on the same principle, inviting active CSP groups to submit a ranked list (up to 100 entries in the 6th blind test and 3 entries previously) of suggested solid structures for unpublished organic structures in different categories. The participating groups are only informed of the chemical structure of the molecules, and the categories vary by the number of atoms, number of molecules in the asymmetric cell and flexibility of the molecule.

There is a great variety in CSP methods with the $6^{th}$ blind test including 25 different submissions but general principles that will apply to all or most methods can be discussed. The starting point of the search is the molecule. While in theory the search can start from atoms, it will usually have a high computational cost and no benefit since the covalent bonds are known. The next stage is then exploring the possible conformations of the molecule. This is more easily achieved for rigid molecules then flexible molecules(*103*), for which the search inherently contains more degrees of freedom. In some methods ab initio calculations of the molecular gas phase determine the initial conformation of the molecule, other methods combine statistical data on conformations from known structures (*104*) or more complex methods can be employed (*105–107*)

After the conformation(s) have been determined, the next stage is to generate candidate crystal structures. Two important parameters to consider are the space group and the number of molecules in the asymmetric unit. There are 230 space groups to explore, however, most organic molecules will crystallise in a rather small subset of space groups. If the unit-cell has only one molecule in the asymmetric unit ($Z' = 1$), then all molecules will be related to each other by symmetry operation, this will significantly simplify the search. Unfortunately, assuming $Z' = 1$ is not a good approximation, since many organic crystal structure have more than a single

molecule in the asymmetric unit ($Z' > 1$), and the different molecules contained in the asymmetric unit are not related to each other by symmetry. The majority of submissions to the $6^{th}$ blind test employed random or quasi-random searches to find candidate crystal structures. Other methods include Monte Carlo simulations, parallel tempering, systematic grid searches, evolutionary and genetic algorithms and shape matching based on known structures in the CSD.

CSP methods will often produce many candidate structures, each corresponding to a local minima in the energy landscape of the investigated compound. The last stage is therefore deciding which of the candidate structures will be realised and which will be the stable form. This is done by a thermodynamic ranking on the candidate structures, assuming the most thermodynamically stable candidate structure will correspond to the experimentally found structure. This assumption is known to not always be correct, yet it is a systematic approach for choosing the most likely structures to form. Ranking is a critical stage since often polymorphs are within a small energy difference of each other, though exceptions with an energy difference of 6-8 kJ $mol^{-1}$ can be found (*108*). The energy ranking, based on 0 K energies, can be determined by DFT with corrections to include the Van der Waals interactions, as was performed in most submissions to the 6th blind test, but other methods can be used, mostly to reduce the computational cost of the energy calculations (*109–111*).

Another important point to consider is the extent of the search. Similar to the experimental solid form screening, one can never know with full certainty that the most thermodynamically stable structure was found, yet the search must cease at some point or another. CSP methods are very computationally expensive. The number of CPU hours used to generate the submissions to the 6h blind tests can be found in the supporting information of the blind test report (*102*), with one entry reported to have used 30,000,000 CPU hours for a single target molecule.

The CSP prediction used in this method work is the Ab initio Random Structure Searching (AIRSS) method (*112*). AIRSS takes the molecular formulation as a starting point, with the free parameters being the number of molecules in the asymmetric unit ($Z'$), the number of molecules in the unit cell ($Z$) and a target volume. Additional constraints can be applied such as symmetry elements or space groups, constraints on the unit cell dimensions and angles, and minimal inter-molecular distances between atomic types. The structure generation is random, and structures are rejected if the constraints are not fulfilled. The structures are then geometry optimised with the DFT code, CASTEP, a stage which includes optimisation of the randomly generated structures, and is also responsible for the ranking of the

structures.

## 2.7 NMR Crystallography

Solid-state NMR spectroscopy and diffraction are two powerful techniques for structural investigation of crystalline materials that are inherently complementary, since they observe the materials on a different scale. XRD probes long range range order, i.e., periodicity, while solid-state NMR probes the ensemble of local chemical environments around the nuclei studied. Combining the two methods can provide stronger structural insights about the studied sample. GIPAW calculations serve as a link between NMR and powder X-ray diffraction (pXRD), and the combined methods from the basis of NMR crystallography (*25*, *113*, *114*).

While single crystal XRD is the most powerful and routine method for structure determination of a solid, it imposes the requirement of growing crystals above a minimal size and with sufficient quality. When a single crystal cannot be obtained, and the compound can only be investigated as powder, pXRD is the next step for structure determination. However, powders are composed of many small crystallites with different orientations, making the structure determination significantly more difficult than with single crystal XRD.

When a structure solution cannot be determined from a single crystal, it can usually be found from pXRD data through Rietveld refinement (*115–119*). The pXRD will detect the heavy elements in the material with great accuracy, but the position of the lighter elements, hydrogen in particular, might be less accurate. The found unit-cell can then be geometry optimised to improve the positions of the hydrogens followed by a calculation of solid-state $^{1}$H and $^{13}$C NMR chemical shifts using the GIPAW approach (*8*, *77*, *120–122*). Comparing the calculated values to the experimentally determined chemical shifts allows a robust independent verification of the structure (*123–129*).

An example for the application of this procedure can be found in Watts *et al.* (*130*). The authors prepared an anhydrous phase of the pharmaceutical material cimetidine hydrochloride. The structure of the anhydrous phase was determined from pXRD data using a Rietveld refinement. A geometry optimisation was carried out using the GIPAW method with TS dispersion correction with a fixed unit cell and maintaining the symmetry of the structure. $^{13}$C magnetic shieldings were then calculated and compared to experimental values. The suggested structure for the anhydrous phase of cimetidine hydrochloride was therefore validated against experimental data from both pXRD and solid state NMR.

Solid-state NMR can provide valuable structural information including: the number of molecules in the asymmetric unit can be assumed from the number of peaks in a high resolution spectrum; the existence of specific interactions and bonds can be detected; the values of certain inter-atomic distances can be measured; and disorder can be discovered.

NMR crystallography has been previously applied to study pharmaceutical co-crystals (*131*, *132*); to investigate the structural effects inducing transformation processes such as dehydration and desolvation (*133–135*); for structure refinements of polycrystalline and macromolecular systems (*136*, *137*); to study amino-acids (*138*); and $^{13}$C NMR in Carbon Nanotubes (*139*) among many other applications. Of specific interest in the context of this work is the analysis of hydrogen bonding through the investigation of chemical shifts (*87*, *140*) and longer range non-covalent interactions (*56*, *141–143*).

Uldry *et al.* (*87*) performed a series of GIPAW calculations combined with experimental solid-state NMR data to investigate weak hydrogen bonding in uracil and 4-Cyano-4′-ethynylbiphenyl. The structures were originally solved form single crystal XRD. $^{1}$H, $^{13}$C, $^{15}$N chemical shifts were investigated by MAS experiments and calculated using GIPAW first principles calculations. The investigation of the bonds was through calculation of artificial yet informative calculations, such as a single uracil molecule, with the atomic positions unchanged from their positions inside the unit cell. Also calculated were the Nucleus-Independent Chemical Shifts (NICS) (*144*), a calculation of the contribution of the neighbouring molecules to the chemical shifts of the investigated molecule. Comparing the calculated contributions from a single molecule, the full crystallographic unit cell, and the NICS allows a careful quantification of hydrogen-bonds and other inter-molecular interactions.

Dudenko *et al.* (*56*) presents an investigation on inter-molecular interactions, specifically hydrogen bonding and $\pi$-$\pi$ interactions in an indomethacin and nicotinamide cocrystal. DFT-D is employed for optimisation of the structure, with two dispersion correction schemes, G06 (*51*) and TS (*50*) employed and compared within the work. Single molecule and NICS calculations are used, and NICS visualisation maps are generated for the visualisation of the effect of aromatic ring currents on the NMR chemical shifts in the indomethacin and nicotinamide cocrystal.

# Statement of Contributions

### Ab-Initio Random Structure Searching of Organic Molecular Solids: Assessment and Validation Against Experimental Data

Initial AIRSS run by others, analysis of candidate structures by thesis author. Diffraction experiments and fitting to data by others, solid-state NMR experiments and analysis by thesis author. Calculation of NMR parameters using GIPAW method by author, including comparison to experimental data. Figures and tables preparation by thesis author, except figures 2-4. Paper writing, thesis author and others.

### Visualising packing interactions in solid-state NMR: Concepts and applications

Theoretical method development by thesis author and others, code written by others, testing and validation of software by author and others, application of method to systems presented in the work by thesis author, data analysis and preparation of all tables and figures by thesis author. Paper writing by thesis author and others.

### An NMR Crystallography investigation of Furosemide

All work by thesis author under supervision, paper writing by thesis author and others.

### Visualization and Processing of Computed Solid-State NMR parameters: MagresView and MagresPython

Thesis author contributed to the development and testing of the Magresview software, in particular to the 2D spectral plotting feature. Thesis author prepared figures 2-6 and contributed to the paper writing.

# Chapter 7

# Summary and Outlook

The work presented in this thesis has both applied and further developed the field of NMR Crystallography. It begins with evaluating the use of the CSP method AIRSS to find polymorphs of an organic solid. While many potential structures can be generated, one significant challenge is identify a specific polymorph based on its experimental signature. For the specific case of mABA, it was shown that out of 600 generated structures the lowest energy structures could be reduced to distinct forms: form III, the most stable form that was indeed found as the lowest energy candidate structures; form IV with a rotated $NH_3$ group and metastable variants; and three structures which do not match any of the known forms. The candidate structures were then compared to experimental solid-state NMR and pXRD data for a successful verification of both forms III and IV. Clearly the use of solid-state NMR in NMR Crystallography relies on the sensitivity of NMR parameters to packing effects in the solid state. Chapter 4 presents a method for decomposing the magnetic shielding into contributions from different origins: local and non-local; hydrogen bonds and long-range packing interactions. Visualisation of the long-range contributions to the NICS is presented in the form of Magnetic Shielding Contribution Field maps. The full analysis is performed on three compounds previously studied for their interesting intermolecular interactions. Chapter 5 presents an NMR crystallography analysis of the API furosemide. Solid state NMR MAS experiments for the $^{1}H$ and $^{13}C$ nuclei were performed on form I of furosemide and assignments were made using GIPAW calculated NMR parameters. The theoretical analysis presented in Chapter 4 was then applied to all three known polymorphs of furosemide. Chapter 6 presents the Magres file format for calculated NMR parameters, and the processing and visualisation tools MagresView and MagresPython.

As demonstrated by the results presented in Chapter 3, CSP and NMR

crystallography have potential to be used in combination in order to solve unsolved structures. While other CSP methods would have also likely found form III and IV of mABA, this work represents an important proof-of-principle study for AIRSS.

The advantage of solid-state NMR when attempting to solve an unknown structure is that an NMR spectrum can provide vital structural information, for example, the number of molecules in the asymmetric unit, prior to the CSP search. However, if we were to attempt to solve an unsolved structure purely from a combination of CSP and solid-state NMR, the structure validation would be difficult. The known discrepancy of the ab-initio calculations as compared to experiment are about 1% of the chemical shift range, making it unlikely that only one structure will fit the solid-state NMR spectra. PXRD, on the other hand, is over selective: a slight variation in the unit cell parameters can make the difference between a successful refinement to the pXRD data and an unsuccessful one. When working with pXRD and solid-state NMR data for an unsolved structure, usually the pXRD data is used to generate a trial structure. The trial structure is then refined and used as a basis for the calculation of NMR parameters. The calculated NMR data is then compared to the experimental NMR data for validation. When the structure cannot be solved via diffraction methods, CSP is employed. In such a case, solid-state NMR is used to gain preliminary structural data, and both techniques are used for validation, with pXRD being the main one.

The work-flow for linking CSP to experiment, however, is still far from being established, and as well as highlighting the potential of using CSP for structure solutions, this work attempts to emphasise the possible pitfalls of the procedure. Perhaps the most significant is the generation of many candidate structures, with the difficulty of predicting which are realisable in reality. In particular, in the CCDC blind tests, a list of 100 ranked structures is allowed for submission, and the search is for the most stable form. In reality, an unsolved structure isn't necessarily the most thermodynamically stable one. A key point, therefore, is to suggest a sensible energetic margin for potentially realisable structures, in this work we used the "safe margin" of 7.2 kJ $mol^{-1}$ following the work of Nyman and Day (*108*), in practice, most polymorphs will be within a much smaller energetic margin (*145*). After reducing the number of candidate structures to a manageable number, identifying the correct structure needs to be carried out by comparison to available experimental data. The challenges presented at this stage can by simplified in claiming that, while the combination of ab initio calculations and solid-state NMR isn't selective enough, due to the discrepancy of calculation compared to experiment stated above, while validation via pXRD can only be done with very accurate unit cell parameters,

which is often a weakness of DFT calculations. Keeping in mind that a structure that is yet to be solved is likely to have a difficulty to fit experimental data, this adds to the already substantial challenge.

As seen in chapter 4, the sensitivity of the magnetic shielding to intermolecular interactions is most noticeable in the chemical shifts of protons. This is because the effect of non-local interactions on chemical shifts is essentially independent of the atomic type, and as protons have the smallest range of chemical shifts, they are most greatly affected. This stresses the importance of acquiring a well-resolved $^{1}$H spectrum, achievable using very fast MAS.

Bridging the gap between CSP and experimental data should be approached from both directions. As many structural insights from the available experimental data should be incorporated into the search in the earliest stage possible. From the other side, a comprehensive theoretical analysis should be performed on the candidate structures within the realisable energetic range. Efforts should be made to identify structural differences that will be detectable experimentally. For example, GIPAW calculated NMR parameters should be compared between candidate structures in order to identify the most notable differences, thus increasing the likelihood of bringing the search closer to a conclusion. Though requiring different expertise, best results are to be expected when the theoretical work is done in tandem with the collection of experimental data.

The ability to turn solid-state NMR data collected experimentally into structural insights will improve alongside our understanding of how packing interactions affect the chemical shift. Even if the unit cell is unknown, molecular conformation analysis can be done, as described in the $6^{th}$ blind test (*102*), arriving at predicted molecular conformations of a single molecular unit. The expected chemical shift for a single molecule could then be calculated and compared to experimental chemical shifts acquired from a high resolution $^{1}$H NMR spectrum and the deviations from the predicted single molecule shifts could be analysed. This could help identify crucial structural motifs and aid the structure solution. For known structures, the analysis can go further, to an in-depth analysis of the packing arrangements of the solid. The core of the DFT method is in solving the electronic density for the ground state. This gives us the ability to try and understand the relationship between the packing interactions and the electronic structure. This is important for better understanding of the relationship between the packing arrangement and the key macroscopic properties presented by the crystal, namely, for pharmaceuticals, solubility and stability.

The tools developed in the course of this work are aimed at serving exper-

imental scientists and make the routinely used GIPAW calculations into a more powerful tool. The intersection between theory and experiment is invaluable to NMR crystallography, and a good understanding of both aspects will promote better scientific research. Emphasis has been placed on making the tools user-friendly and intuitive, particularly for scientists without a strong computational background.

**Future work**

Perhaps the most important contribution of this work to the field on NMR crystallography is that it explores and aims to enable new applications of NMR crystallography through better synergy of experimental data with the ever growing computational capabilities. This work can be used as the basis for developing a tool suite to better integrate CSP methods in general, and AIRSS in particular, with experimental pXRD and solid-state NMR data, maximising the chances of solving previously unsolved structures. Furthermore, the tools developed in this work, and in chapter 4 specifically, enable in depth investigation of inter-molecular interactions in molecular crystals. This can be used for the investigation of polymorphic materials, to better understand how the crystal packing affects the different properties of exhibited by different polymorphs of the same compound.

# Bibliography


1. S. Sturniolo *et al.*, *Solid State Nucl. Magn. Reson.* **78**, 64–70 (2016).
2. M. Zilka *et al.*, *Phys. Chem. Chem. Phys.* **19**, 25949–25960 (2017).
3. M. Zilka, S. Sturniolo, S. P. Brown, J. R. Yates, *J. Chem. Phys.* **147**, 144203 (2017).
4. E. M. Purcell, H. C. Torrey, R. V Pound, *Phys. Rev.* **69**, 37 (1946).
5. F. Bloch, W. W. Hansen, M Packard, *Phys. Rev.* **70**, 474–485 (1946).
6. E. R. Andrew, A. Bradbury, R. G. Eades, *Nature* **182**, 1659–1659 (1958).
7. C. J. Pickard, F. Mauri, *Phys. Rev. B* **63**, 245101 (2001).
8. J. R. Yates, C. J. Pickard, F. Mauri, *Phys. Rev. B* **76**, 24401 (2007).
9. C. J. Pickard, R. Needs, *Nat. Mater.* **9**, 624–627 (2010).
10. C. J. Pickard, M. Martinez-Canales, R. J. Needs, *Phys. Rev. B* **85**, 214114 (2012).
11. M. Martinez-Canales, C. J. Pickard, R. J. Needs, *Phys. Rev. Lett.* **108**, 045704 (2012).
12. S. Zhang, H. F. Wilson, K. P. Driver, B. Militzer, *Phys. Rev. B* **87**, 024112 (2013).
13. M. Mayo, K. J. Griffith, C. J. Pickard, A. J. Morris, *Chem. Mater.* **28**, 2011–2021 (2016).
14. J. M. Stratford *et al.*, *J. Am. Chem. Soc.* **139**, 7273–7286 (2017).
15. M. Adobes-Vidal *et al.*, *Cryst. Growth Des.* **16**, 4421–4429 (2016).
16. N. R. Goud *et al.*, *J. Pharm. Sci.* **101**, 664–680 (2012).
17. C. Garnero, A. K. Chattah, M. Longhi, *J. Pharm. Biomed. Anal.* **95**, 139–145 (2014).

*18.* W. C McCrone, *In "Physics and Chemistry of the Organic Solid State"; Fox, D., Labes, M. M., Weissberger, A., Eds. Wiley Interscience: New York* **2**, 725–767 (1965).

*19.* E. H. Lee, *Asian J. Pharm. Sci.* **9**, 163–175 (2014).

*20.* A. Llinás, J. M. Goodman, *Drug Discovery Today* **13**, 198–210 (2008).

*21.* J. Bernstein, *Polymorphism in Molecular Crystals* (Oxford University Press, 2007).

*22.* T. L. Threlfall, *Analyst* **120**, 2435–2460 (1995).

*23.* A. J. Cruz-Cabeza, S. M. Reutzel-Edens, J. Bernstein, *Chem. Soc. Rev.* **44**, 8619–8635 (2015).

*24.* J. Haleblian, W. McCrone, *J. Pharm. Sci.* **58**, 911–929 (1969).

*25.* R. K. Harris, *Solid State Sci.* **6**, 1025–1037 (2004).

*26.* J. C. Chan, M. Bertmer, H. Eckert, *J. Am. Chem. Soc.* **121**, 5238–5248 (1999).

*27.* D. C. Apperley *et al.*, *Magn. Reson. Chem.* **43**, 881–892 (2005).

*28.* T. N. Pham *et al.*, *Mol. Pharm.* **7**, 1667–1691 (2010).

*29.* J. A. Ripmeester, *Chem. Phys. Lett.* **74**, 536–538 (1980).

*30.* S. P. Brown, *Solid State Nucl. Magn. Reson.* **41**, 1–27 (2012).

*31.* A. J. Aguiar, J. Krc, A. W. Kinkel, J. C. Samyn, *J. Pharm. Sci.* **56**, 847–53 (1967).

*32.* M. Draguet-Brughmans, R. Bouche, J. Flandre, A. Vandenbulke, *Pharm. Acta Helv.* **54** (1979).

*33.* A. V. Thomas *et al.*, *Org. Process Res. Dev.* **1**, 294–299 (1997).

*34.* H. G. Brittain, *Polymorphism in pharmaceutical solids* (Informa Healthcare, 2009), p. 640.

*35.* H. G. Brittain, *J. Pharm. Sci.* **101**, 464–484 (2012).

*36.* J. Chen, B. Sarma, J. M. Evans, A. S. Myerson, *Cryst. Growth Des.* **11**, 887–895 (2011).

*37.* S. L. Morissette *et al.*, *Adv. Drug Deliv. Rev.* **56**, 275–300 (2004).

*38.* V. V. Chernyshev, *Russ. Chem. Bull.* **50**, 2273–2292 (2001).

*39.* S. Perez-Lloret, M. V. Rey, P. L. Ratti, O. Rascol, *Fundam. Clin. Pharmacol.* **27**, 81–95 (2013).

*40.* P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864–B871 (1964).

*41.* W. Kohn, L. J. Sham, *Phys. Rev.* **140**, A1133–A1138 (1965).

*42.* D. M. Ceperley, B. J. Alder, *Phys. Rev. Lett.* **45**, 566–569 (1980).

*43.* J. P. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048–5079 (1981).

*44.* J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

*45.* O. A. Vydrov, T. Van Voorhis, *Phys. Rev. Lett.* **103**, 063004 (2009).

*46.* H. Rydberg *et al.*, *Phys. Rev. Lett.* **91**, 126402 (2003).

*47.* J. F. Dobson, B. P. Dinte, *Phys. Rev. Lett.* **76**, 1780–1783 (1996).

*48.* J. Klimeš, A. Michaelides, *J. Chem. Phys.* **137**, 120901 (2012).

*49.* A. R. Ferreira, E. Küçükbenli, A. A. Leitão, S. de Gironcoli, *Phys. Rev. B* **84**, 235119 (2011).

*50.* A. Tkatchenko, M. Scheffler, *Phys. Rev. Lett.* **102**, 73005 (2009).

*51.* S. Grimme, *J. Comput. Chem.* **27**, 1787–1799 (2006).

*52.* F. Ortmann, F. Bechstedt, W. G. Schmidt, *Phys. Rev. B* **73**, 205101 (2006).

*53.* P. Jurečka, J. Černý, P. Hobza, D. R. Salahub, *J. Comput. Chem.* **28**, 555–569 (2007).

*54.* I. A. Fedorov, Y. N. Zhuravlev, V. P. Berveno, *Phys. Chem. Chem. Phys.* **13**, 5679 (2011).

*55.* S. Hedström, P. Persson, *J. Phys. Chem. C* **116**, 26700–26706 (2012).

*56.* D. V. Dudenko, J. R. Yates, K. D. M. Harris, S. P. Brown, *CrystEngComm* **15**, 8797 (2013).

*57.* M. H. Levitt, *Spin Dynamics* (Wiley, Chichester, 2001).

*58.* M. J. Duer, *Introduction to Solid-State NMR Spectroscopy* (Blackwell, Oxford, Oxford, 2004).

*59.* P. J. Hore, J. A. Jones, S. Wimperis, *NMR: The Toolkit* (Oxford University Press, Oxford, 2000).

*60.* M. J. Duer, *Principles and Applications* (Blackwell, Oxford, Oxford, 2007).

*61.* M. J. Duer, in *Solid-State NMR Spectroscopy Principles and Applications* (Blackwell Science Ltd, 2007), pp. 1–72.

*62.* J. H. Keeler, in *Multinuclear Magnetic Resonance in Liquids and Solids — Chemical Applications*, ed. by P. Granger, R. K. Harris (Springer Netherlands, 1990), pp. 201–238.

*63.* S. R. Hartmann, E. L. Hahn, *Phys. Rev.* **128**, 2042–2053 (1962).

*64.* G. Metz, X. L. Wu, S. O. Smith, *J. Magn. Reson., Series A* **110**, 219–227 (1994).

*65.* G. A. Morris, R. Freeman, *J. Am. Chem. Soc.* **101**, 760–762 (1979).

*66.* D. P. Burum, R. R. Ernst, *J. Magn. Reson.* **39**, 163–168 (1980).

*67.* B. Elena, A. Lesage, S. Steuernagel, A. Böckmann, L. Emsley, *J. Am. Chem. Soc.* **127**, 17296–17302 (2005).

*68.* D. Sakellariou, A. Lesage, P. Hodgkinson, L. Emsley, *Chem. Phys. Lett.* **319**, 253–260 (2000).

*69.* A. Lesage *et al.*, *J. Magn. Reson.* **163**, 105–113 (2003).

*70.* D. P. Burum, M. Linder, R. R. Ernst, *J. Magn. Reson.* **44**, 173–188 (1981).

*71.* B. Elena, G. de Paëpe, L. Emsley, *Chem. Phys. Lett.* **398**, 532–538 (2004).

*72.* W. Sommer, J. Gottwald, D. Demco, H. Spiess, *J. Magn. Reson., Series A* **113**, 131–134 (1995).

*73.* I. Schnell, A. Lupulescu, S. Hafner, D. E. Demco, H. W. Spiess, *J. Magn. Reson.* **133**, 61–9 (1998).

*74.* S. P. Brown, *Prog. Nucl. Magn. Reson. Spectrosc.* **50**, 199–251 (2007).

*75.* I. Schnell, H. W. Spiess, *J. Magn. Reson.* **151**, 153–227 (2007).

*76.* S. Hafner, H. W. Spiess, *Concepts Magn. Reson.* **10**, 99–128 (1998).

*77.* C. Bonhomme *et al.*, *Chem. Rev.* **112**, 5733–5779 (2012).

*78.* D. Vanderbilt, *Phys Rev B* **41**, 7892–7895 (1990).

*79.* K. Lejaeghere *et al.*, *Science* **351** (2016).

*80.* J. R. Yates, C. J. Pickard, in *NMR Crystallography*, ed. by R. K. Harris, R. E. Wasylishen, M. J. Duer (Wiley, 2009), chap. 3, pp. 23–40.

*81.* F. Mauri, B. G. Pfrommer, S. G. Louie, *Phys. Rev. Lett.* **77**, 5300–5303 (1996).

*82.* F. Mauri, S. G. Louie, *Phys. Rev. Lett.* **76**, 4246–4249 (1996).

*83.* C. G. van de Walle, P. E. Blöchl, *Phys. Rev. B* **47**, 4244–4255 (1993).

*84.* S. J. Clark *et al.*, *Z. Kristall.* **220**, 567–570 (2005).

*85.* P. Giannozzi *et al.*, *J. Phys. Condens. Matter* **21**, 395502 (2009).

*86.* C. Gervais *et al.*, *Magn. Reson. Chem.* **42**, 445–452 (2004).

*87.* A. C. Uldry *et al.*, *J. Am. Chem. Soc.* **130**, 945–954 (2008).

*88.* G. H. Penner, R. Webber, L. A. O'Dell, *Can. J. Chem.* **89**, 1036–1046 (2011).

*89.* R. K. Harris *et al.*, *Org. Process Res. Dev.* **9**, 902–910 (2005).

*90.* R. K. Harris, S. A. Joyce, C. J. Pickard, S. Cadars, L. Emsley, *Phys. Chem. Chem. Phys.* **8**, 137–143 (2006).

*91.* R. K. Harris *et al.*, *Cryst. Growth Des.* **8**, 80–90 (2008).

*92.* M. D. Esrafili, *Can. J. Chem.* **89**, 1410–1418 (2011).

*93.* A. Abraham, D. C. Apperley, T. Gelbrich, R. K. Harris, U. J. Griesser, *Can. J. Chem.* **89**, 770–778 (2011).

*94.* H. J. Monkhorst, J. D. Pack, *Phys. Rev. B* **13**, 5188–5192 (12 1976).

*95.* A. Zheng, S. B. Liu, F. Deng, *J. Phys. Chem. C* **113**, 15018–15023 (2009).

*96.* J. M. Griffin, J. R. Yates, A. J. Berry, S. Wimperis, S. E. Ashbrook, *J. Am. Chem. Soc.* **132**, 15651–15660 (2010).

*97.* M. Profeta, M. Benoit, F. Mauri, C. J. Pickard, *J. Am. Chem. Soc.* **126**, 12628–12635 (2004).

*98.* S. L. Price, D. E. Braun, S. M. Reutzel-Edens, *Chem. Commun.* **52**, 7065–7077 (2016).

*99.* M. A. Neumann, F. J. Leusen, J. Kendrick, *Angew. Chem. Int. Ed.* **47**, 2427–2430 (2008).

*100.* G. M. Day, *Crystallogr. Rev.* **17**, 3–52 (2011).

*101.* J. Kendrick, F. J. Leusen, M. A. Neumann, J. van de Streek, *Chem. Eur. J.* **17**, 10736–10744 (2011).

*102.* A. M. Reilly *et al.*, *Acta Crystallogr., Sect. B: Struct. Sci* **72**, 439–459 (2016).

*103.* G. M. Day, W. S. Motherwell, W Jones, *PCCP* **9**, 1693–1704 (2007).

*104.* I. J. Bruno *et al.*, *J. Chem. Inform. Comput. Sci.* **44**, 2133–2144 (2004).

*105.* I. Kolossváry, W. C. Guida, *J. Am. Chem. Soc.* **118**, 5011–5019 (1996).

*106.* H. Gotō, E. Ōsawa, *J. Am. Chem. Soc.* **111**, 8950–8951 (1989).

*107.* H. Gotō, E. Ōsawa, *Perkin Trans. 2*, 187–198 (1993).

*108.* J. Nyman, G. M. Day, *CrystEngComm* **17**, 5154–5165 (2015).

*109.* M. A. Neumann, *J. Phys. Chem. B* **112**, 9810–9829 (2008).

*110.* J. van de Streek, M. A. Neumann, *CrystEngComm* **13**, 7135–7142 (2011).

*111.* M. Neumann, J Van De Streek, F. Fabbiani, P Hidber, O Grassmann, *Nat. Commun.* **6** (2015).

*112.* C. J. Pickard, R. J. Needs, *J. Phys. Condens. Matter* **23**, 053201 (2011).

*113.* B. Elena, G. Pintacuda, N. Mifsud, L. Emsley, *J. Am. Chem. Soc.* **128**, 9555–9560 (2006).

*114.* R. K. Harris, R. E. Wasylishen, M. J. Duer, *NMR Crystallography* (Wiley, 2009), p. 504.

*115.* P. Lightfoot *et al.*, *Chem. Comm.* **47**, 1012 (1992).

*116.* K. D. M. Harris, M. Tremayne, P. Lightfoot, P. G. Bruce, *J. Am. Chem. Soc.* **116**, 3543–3547 (1994).

*117.* K. D. M. Harris, M. Tremayne, B. M. Kariuki, *Angew. Chem. Int. Ed.* **40**, 1626–1651 (2001).

*118.* M. Tremayne, *Philos. Trans. R. Soc. London, Ser. A* **362** (2004).

*119.* W. I. F. David *et al.*, *Acta Crystallogr., Sect. A: Found. Crystallogr.* **64**, 52–64 (2008).

*120.* C. J. Pickard, F. Mauri, *Phys. Rev. B* **63**, 245101 (2001).

*121.* R. K. Harris, P. Hodgkinson, C. J. Pickard, J. R. Yates, V. Zorin, *Magn. Reson. Chem.* **45**, S174–S186 (2007).

*122.* T. Charpentier, *Solid State Nucl. Magn. Reson.* **40**, 1–20 (2011).

*123.* X. Filip, G. Borodi, C. Filip, *Phys. Chem. Chem. Phys.* **13**, 17978–17986 (2011).

*124.* X. Filip, I.-G. Grosu, M. Miclăuş, C. Filip, *CrystEngComm* **15**, 4131–4142 (2013).

*125.* D. V. Dudenko *et al.*, *J. Phys. Chem. C* **117**, 12258–12265 (2013).

*126.* X. Li *et al.*, *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.* **70**, 784–789 (2014).

*127.* P. Li *et al.*, *CrystEngComm* **16**, 3141–3147 (2014).

*128.* M. Sardo *et al.*, *Solid State Nucl. Magn. Reson.* **65**, 49–63 (2015).

*129.* G. N. M. Reddy, A. Marsh, J. T. Davis, S. Masiero, S. P. Brown, *Cryst. Growth Des.* **15**, 5945–5954 (2015).

*130.* A. E. Watts, K. Maruyoshi, C. E. Hughes, S. P. Brown, K. D. Harris, *Cryst. Growth Des.* **16**, 1798–1804 (2016).

*131.* O. D. Putra *et al.*, *Cryst. Growth Des.* **16**, 5223–5229 (2016).

*132.* D. Luedeker, R. Gossmann, K. Langer, G. Brunklaus, *Cryst. Growth Des.* **16**, 3087–3100 (2016).

*133.* M. K. Dudek, T. Pawlak, P. Paluch, A. Jeziorna, M. J. Potrzebowski, *Cryst. Growth Des.* **16**, 5312–5322 (2016).

*134.* A. K. Chattah *et al.*, *Mol. Pharm.* **12**, 731–741 (2015).

*135.* F. Costantino, P. Sassi, M. Geppi, M. Taddei, *Cryst. Growth Des.* **12**, 5462–5470 (2012).

*136.* J. Brus *et al.*, *Macromolecules* **47**, 6343–6354 (2014).

*137.* P. Matějíček *et al.*, *Macromolecules* **44**, 3847–3855 (2011).

*138.* C. Gervais *et al.*, *J. Phys. Chem. A* **109**, 6960–6969 (2005).

*139.* M. Kibalchenko, M. C. Payne, J. R. Yates, *ACS Nano* **5**, 537–545 (2011).

*140.* J. R. Yates *et al.*, *J. Am. Chem. Soc.* **127**, 10216–10220 (2005).

*141.* J. P. Bradley, S. P. Velaga, O. N. Antzutkin, S. P. Brown, *Cryst. Growth Des.* **11**, 3463–3471 (2011).

*142.* G. M. Reddy *et al.*, *Solid State Nucl. Magn. Reson.* **65**, 41–48 (2015).

*143.* E. Carignani, S. Borsacchi, J. P. Bradley, S. P. Brown, M. Geppi, *J. Phys. Chem. C* **117**, 17731–17740 (2013).

*144.* D. Sebastiani, *ChemPhysChem* **7**, 164–175 (2006).

*145.* S. L. Price, *Chem. Soc. Rev.* **43**, 2098–2111 (2014).